

STIC-Biotech/ChemLib

172599

ME

From: Li, Bao-Qun
Sent: Monday, November 28, 2005 11:55 AM
To: STIC-Biotech/ChemLib
Subject: 10,761,006

Please do the amino acid sequence homology and interference search against the nucleic acid sequence of SEQ ID NO: 1.

Bao Qun Li M.D
TC 1600
Art Unit 1648
Tel. 517-272-0904
REM, 3C18
Rm. 3D24

RECEIVED
NOV 28 2005
STIC/CHEM. DIVISION
(STIC)

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 1, 2005, 23:14:50 ; Search time 706.5 Seconds
(without alignments)
6421.156 Million cell updates/sec

Title: US-10-761-006A-1
Perfect score: 5907
Sequence: 1 CTCACACATTCACCAAG.....CCTCAGCCACGAGTGGA 3215

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgm2.1/USFO.spool_p/US10761006/runat_01122005_113952_26869/app_query.fasta_1.3399
-DB=UniProt -QWTF=fastcan -SUFFX=n2p.rup -MINMATCH=0.1 -LOOPCL=0 -LOOPTXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10761006.qcgm 1.1199 @runat_01122005_113952_26869 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEXTIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt 05.80.*
1: uniprot.spot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	4130	69.9	843	2	Q69616_HPBVO	Q69616 hepatitis b
2	3745.5	63.4	832	2	Q67907_HPBVO	Q67907 hepatitis b
3	3693.5	62.5	842	2	Q69590_HPBVO	Q69590 hepatitis b
4	2824	47.8	843	2	Q7TDR5_HPBVO	Q7TDR5 hepatitis b
5	2822	47.8	843	2	Q4FD83_HPBVO	Q4FD83 hepatitis b
6	2810	47.6	840	2	Q69026_HPBVO	Q69026 hepatitis b
7	2808	47.5	843	2	Q69049_HPBVO	Q69049 hepatitis b
8	2807	47.5	843	1	DROL_HPBVR	P03157 hepatitis b
9	2806	47.5	843	2	Q97975_HPBVO	Q97975 hepatitis b
10	2806	47.5	843	2	Q5KR39_HPBVO	Q5KR39 hepatitis b
11	2804	47.5	843	2	Q7TDR3_HPBVO	Q7TDR3 hepatitis b
12	2803	47.5	843	2	Q68RQ2_HPBVO	Q68RQ2 hepatitis b
13	2803	47.5	843	2	Q4FD65_HPBVO	Q4FD65 hepatitis b
14	2802	47.4	838	2	Q81165_HPBVO	Q81165 hepatitis b
15	2802	47.4	843	2	Q42041_HPBVO	Q42041 hepatitis b
16	2802	47.4	843	2	Q5KR23_HPBVO	Q5KR23 hepatitis b

17	2801	47.4	843	2	Q80MR4_HPBVO	Q80MR4 hepatitis b
18	2800	47.4	843	2	Q5DVZ8_HPBVO	Q5DVZ8 hepatitis b
19	2799	47.4	843	2	Q5KR35_HPBVO	Q5KR35 hepatitis b
20	2799	47.4	843	2	Q4FD61_HPBVO	Q4FD61 hepatitis b
21	2798	47.4	843	2	Q91549_HPBVO	Q91549 hepatitis b
22	2798	47.4	843	2	Q80H07_HPBVO	Q80H07 hepatitis b
23	2798	47.4	843	2	Q5KR19_HPBVO	Q5KR19 hepatitis b
24	2797	47.4	843	2	Q6YLM0_HPBVO	Q6YLM0 hepatitis b
25	2796	47.3	843	2	Q80MQ8_HPBVO	Q80MQ8 hepatitis b
26	2796	47.3	843	2	Q8JVC9_HPBVO	Q8JVC9 hepatitis b
27	2795	47.3	837	2	Q5SDK1_HPBVO	Q5SDK1 hepatitis b
28	2794	47.3	843	2	Q9QNS4_HPBVO	Q9QNS4 hepatitis b
29	2794	47.3	843	2	Q9VZB5_HPBVO	Q9VZB5 hepatitis b
30	2794	47.3	843	2	Q8BC17_HPBVO	Q8BC17 hepatitis b
31	2792	47.3	843	2	Q91547_HPBVO	Q91547 hepatitis b
32	2792	47.3	843	2	Q99217_HPBVO	Q99217 hepatitis b
33	2792	47.3	843	2	Q9QM17_HPBVO	Q9QM17 hepatitis b
34	2792	47.3	843	2	Q9QMJ8_HPBVO	Q9QMJ8 hepatitis b
35	2791.5	47.3	844	2	Q4FD77_HPBVO	Q4FD77 hepatitis b
36	2791	47.2	843	2	Q97976_HPBVO	Q97976 hepatitis b
37	2791	47.2	843	2	Q39671_HPBVO	Q39671 hepatitis b
38	2791	47.2	843	2	Q81134_HPBVO	Q81134 hepatitis b
39	2791	47.2	843	2	Q8VIM4_HPBVO	Q8VIM4 hepatitis b
40	2790	47.2	843	2	Q91541_HPBVO	Q91541 hepatitis b
41	2790	47.2	843	2	Q5SDK5_HPBVO	Q5SDK5 hepatitis b
42	2790	47.2	843	2	Q5KR43_HPBVO	Q5KR43 hepatitis b
43	2790	47.2	843	2	Q5R2L9_HPBVO	Q5R2L9 hepatitis b
44	2790	47.2	843	2	Q9QMK4_HPBVO	Q9QMK4 hepatitis b
45	2789	47.2	843	2	Q9PWY8_HPBVO	Q9PWY8 hepatitis b

ALIGNMENTS

RESULT 1
Q69616_HPBVO
ID Q69616_HPBVO PRELIMINARY; PRT; 843 AA.
AC Q69616;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE Polymerase (Fragment).
GN NameP;
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Genotype c;
RX MEDLINE=94120723; PubMed=8291231;
RA Norder H., Courouce A.M., Magnus L.O.;
RT "Complete genomes, phylogenetic relatedness, and structural proteins of six strains of the hepatitis B virus, four of which represent two new genotypes.";
RL Virology 198:489-503(1994).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Genotype c;
RA Norder H.M.L.;
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; X75656; CAAS3338.1; -; Genomic DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001462; DNaPol_viral_C.
DR InterPro; IPR000201; DNaPol_viral_N.
DR InterPro; IPR000477; RVise.
DR Pfam; PF00336; DNA_pol_viral_C; 1.
DR Pfam; PF00242; DNA_pol_viral_N; 1.

DR Pfam; PF00078; RVT 1; 1.
 DR ProDom; PD000814; DNapol_viral_C; 1.
 DR PROSITE; P550878; RT_POL; 1.
 FT CHAIN 541 588 core antigen.
 FT NON_TER 1 1
 FT NON_TER 843 843
 SQ SEQUENCE 843 AA; 94603 MW; 934EEBCIF4E235D0 CRC64;

Alignment Scores:
 Pred. No.: 1.55e-299 Length: 843
 Score: 4130.00 Matches: 802
 Percent Similarity: 76.26% Conservative: 17
 Best Local Similarity: 74.67% Mismatches: 24
 Query Match: 69.92% Indels: 232
 DB: 2 Gaps: 1

US-10-761-006A-1 (1-3215) x Q69616_HPBVO (1-843)

QY 1 CTCACAAATCCACCAAGCTCTGTAGATCCAGGGGTGAGGGCGCTATATTTCTCTGC 60
 Db 1 LeuHisAsnIleProThrSerSerAlaArgSerGlnSerGluGlyProIlePheSerCys 20
 QY 61 TGGTGGCTCCAGTTCGGGACAGTAACCCGTGTCGACTACTGCTCTCCCATATCGTC 120
 Db 21 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerHisPheVal 40
 QY 121 AATCTTCTCGAGACTCGGGGACCTCGACCGAACATGGAGAACACAAATCAGGATTCCT 180
 Db 41 AsnLeuLeuGluAspTrpGlyProCysAsnGluHisGlyLHisAsnIleArgIlePro 60
 QY 181 AGAACCCCTGCTGTTTACAGCGGGGTTTTCTCGTTTGAACAAGAACTCTCACAAATACC 240
 Db 61 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 80
 QY 241 GCAGAGCTAGACTCTG-GTGGACTCTCTCAATTTCTAGGGGAGGACCCACGCTGTC 299
 Db 81 ThrGluSerArgLeuGlyValAspPheSerGlnPheSerArgGlySerThrLysVal-Se 100
 QY 300 CTGGCCAAATTCGCAAGTCCCAACCTCCAATCAGTCCCAACCTCTGCTCTCCCAATTT 359
 Db 100 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeu 120
 QY 360 GTCTGGCTATCGCTGGATGTGTCTCGCGGTTTTATCATATTCCTTCTTCATCTGCTGC 419
 Db 120 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaAl 140
 QY 420 TATGCTCATCTCTTGTGTGTTCTCTCGACTACCAAGGTATGTGCGCGTTTGTCTCTC 479
 Db 140 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 160
 QY 480 TACTTCCAGGAACATCAACACAGCAGCGGGCCATGCAAGACCTGACAGCTCTCTGCTC 539
 Db 160 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 180
 QY 540 AAGAAACTCTAGTTTCCCTCTGTGTGTCTGTATACAAAACCTTCGGACGGAAACTGCACTT 599
 Db 180 rArgAsnLeuTyrValSerLeuMetLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 200
 QY 600 GTATTCCCATCCATCATCTGGGCTTTCGCAAGATTCCTATGGGAGTGGGCTCAGTCC 659
 Db 200 uTyrSerHisProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 220
 QY 660 GTTTCCTCGGCTCAGTTACTAGTCCCATTTGTTTCAGTGGTTCTGAGGGCTTTCCCCCA 719
 Db 220 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi 240
 QY 720 CTGTTTGGCTTTTCAGTTATATGATGATGTGGTATTTGGGGGCGAAGTCTGTACAACTCT 779
 Db 240 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 260
 QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCC 839
 Db 260 uGluSerLeuTyrThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 280

QY 840 TAATAAAACCAACAGCTTGGGCTACTCCCTTAACCTTATCGGATATGTAATTGGAAGTTG 899
 Db 280 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 300
 QY 900 GGGTACTTTACCGCAGGACATATTTGTACTATAAACTCAAGCNAATGTTTTCGAAAACTGCC 959
 Db 300 pGlyThrLeuProGlnGluHisIleValLeuLysIleLysGlnCysPheArgLysLeuPr 320
 QY 960 TGTAAATAGACCTATTGATGTGAAAGATATGTCAAAGAATTGTGGGCTCTTTTGGGCTTTGC 1019
 Db 320 ovalAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 340
 QY 1020 TGCCCCCTTTTACAAATGTGGCTATCTCTGCCCTTGTATGCTTTTATATGATGATATCAATC 1079
 Db 340 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnAl 360
 QY 1080 TAAAGAGGCTTTACTTTCTCGCCCAACTTACAAGCCCTTTCTGTGTAACAATATCTGAA 1139
 Db 360 aLysGlnAlaPheThrPheSerSerThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 380
 QY 1140 CCTTTTACCCCTGTCGCCCAACGCTCGGCTCTCTGCCAAGTGTGTGCTGACGCAACCCC 1199
 Db 380 nLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 400
 QY 1200 CACTCGATGGGGCTTTGGCCATAGGCATCAGCCGATGGCTGGAACTTTCTGGCTCTCT 1259
 Db 400 oThrGlyTrpGlyLeuAlaIleGlyAsnGlnArgMetArgGlyThrPheValAlaProLe 420
 QY 1260 GCCGATCCATATCGGGAACCTCTAGACAGCTTTGTTTGTCTGCACGCCGCTGTGGAGCAA 1319
 Db 420 uProIleHisThrAlaGluLeuLeuAlaCysPheAlaArgSerArgSerGlyAlaLys 440
 QY 1320 ACTTATCGAACCGCAACTCTGTTGTCTCTCTCGGAAATACACTCTTCTTCCATGGCT 1379
 Db 440 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 460
 QY 1380 GCTAGGCTGTCTGCAACTGGATCTCTGGCGGAGCTCTTTGTCTACGTCCCTCGGC 1439
 Db 460 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 480
 QY 1440 GCTGAATCCCGGAGACGCCGTCTCGGGGCGGTTTGGGGCTCTACCGTCCCTCTTCA 1499
 Db 480 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyIleTyrArgProLeuLeuAr 500
 QY 1500 TCTGCGCTTTCGGCGCAGCACGGGCGCACCTCTCTTACCGCGTCTCCCGTATGTGCC 1559
 Db 500 gLeuArgPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 520
 QY 1560 TTCTCATCTGCGGACCGGTGTGCACTTCGCTTCACTCTGACGTCGATGGAGACCACC 1619
 Db 520 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 540
 QY 1620 GTGAACGCAACGCGAGGTCTTGCCCAAGGTCTTATATAAGAGACTCTTGGACTCTCAGCA 1679
 Db 540 o----- 540
 QY 1680 ATGTCAAACGACCGACTTTGAGGCATACCTTCAAAGACTGTGTGTTTAAAGACTGGGAG 1739
 Db 540 ----- 540
 QY 1740 TTGGGGGAGAGATTAGGTTAAAGATTATGTACTAGGAGGCTGTAGGCATAAATTTGGTC 1799
 Db 540 ----- 540
 QY 1800 TGTTCACGACGACCATGCAACTTTTCTCTCTGCTGCTAATCATCTCATGTTTCATGCTTA 1859
 Db 540 ----- 540
 QY 1860 CTGTTCAAGCCCTCAAGCTGTGCTTGGGTGGCTTTGGGACATGGACATTGACCCGTATA 1919
 Db 540 ----- 540


```
QY 1920 AAGAAATTGGAGCATCTGCTGAGTACTCTCTTTTGGCCCTTCTGACTCTTTCCGCTCTA 1979
Db 540 -----
QY 1980 TTCGAGATCTCTCGACACCGCCTCTGCTCTGTATCGGAGGCTTAGAGTCTCCGGAAC 2039
Db 540 -----
QY 2040 ATTGTCGCCTCACCATACAGCACTCAGGCAAGCTATTTTGTGTGGGTGGAGTGTGATGA 2099
Db 540 -----
QY 2100 ATCTGGCCACTCGGTGGGAAGTAATTTGGAGATCCAGCATCCAGGAATTAGTAGTCA 2159
Db 540 -----
QY 2160 GCTATGTCAAGTTAATATGGCCCTAACTCAGACAAATATTGTGTTTTCACATTTCCCT 2219
Db 540 -----
QY 2220 GTCTTACTTTTGGAGAGAACTGTCTTGTAGTACTTGGTATCTTTTGGAGTGTGGATTC 2279
Db 540 -----
QY 2280 GCACTCCTACCGTTACAGACCACAAATGCCCTATCTTATCAACTTCCGGAACCTA 2339
Db 541 -----Met--ProLeuSerTyrGlnHisPheArgLysLeu 551
QY 2340 CTGTTGTTTAGACGAGGAGGCTCCCTAGAGAAGAACTCCCTCGCTCGCAGACGAA 2399
Db 552 LeuLeuLeuAspValGluAlaGlyProLeuGluGluLeuProArgLeuAlaAspGlu 571
QY 2400 GGTCTCAATCCCGCGTCGAGAGATCTCAATCTCGGGAATCTCAACGTTAGTATTCCT 2459
Db 572 GlyLeuAsnHisArgValAlaGluAspLeuAsnLeuGlyAsnProAsnValSerIlePro 591
QY 2460 TGGACTCATAAAGTGGGAACTTTACTGGGCTTTTATCTTCTACTGACTGCTTTTAAT 2519
Db 592 TrpThrHisValGlyAsnPheThrGlyLeuTyrSerSerThrValProValPheAsn 611
QY 2520 CCCGAGTGGCAATTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2579
Db 612 ProGluTrpGlnThrProSerPheProAspIleHisLeuGlnGluAspIleIleAsnArg 631
QY 2580 TGTCAACAATATGTGGGCTCTTACAGTTAATGAAAGAAAGAAATTAATTAATATG 2639
Db 632 CysGlnGlnTyrValGlyProLeuThrValAsnGluLysArgLeuLysLeuIleMet 651
QY 2640 CTGCTAGTTTATCTTAACCTTACTAAATATTTGGCTTTAGACAAAGGCATTAACCG 2699
Db 652 ProAlaArgPheTyrProAsnLeuThrLysTyrLeuProLeuAspLysGlyIleLysPro 671
QY 2700 TATTATCTGACATGCGATTAACTATCTTCAAACTAGGCAATTTTACATCTCTG 2759
Db 672 TyrTyrProGluHisAlaValAsnHisTyrPheLysThrArgHisTyrLeuHisThrLeu 691
QY 2760 TGGAGGCTGGCATTTCTATATAAGAGAGAACTACAGCAGCGCTCATTTTGGGTGCTA 2819
Db 692 TrpLysAlaGlyIleLeuTyrLysArgGluThrThrArgSerAlaSerPheCysGlySer 711
QY 2820 CCATATCTTGGGAACAAGAGCTACAGATCGGAGGTGGTCTTCCAAACCTCCGACAAAG 2879
Db 712 ProTyrSerTrpGluGlnGluLeuGlnHisGlyArgLeuValPheGlnThrSerGluArg 731
QY 2880 CATGGGAGCAATCTTGTCTGTCTCCAACTCTCGGATCTTCTCCGATCACAGTTGGA 2939
Db 732 HisGlyAspGluSerPheCysSerGlnSerGlyIleLeuSerArgSerProValGly 751
QY 2940 CCCTCGTTTCGGAGCAACTCAACAATCCAGATGGGACTTCAACCCCAACCAAGGATCA 2999
Db 752 ProCysIleArgSerGlnLeuLysGlnSerArgLeuGlnProGlnGlnGlySer 771
QY 3000 CTGGCCAGGAGCAATCAGGTAGGAGTGGGAGCATTCGGGGCCAGGGTTTCAACCCACCA 3059
```

```
Db 772 MetAlaArgGlySerGlyArgSerGlyArgSerGlySerIleArgAlaArgValHisProThrThr 791
QY 3060 CGCGCGTCTTTTGGGGGAGCCCTCAGGCTCAGGCATATTGACAACTGCGCAGCAGC 3119
Db 792 ArgArgSerPheGlyValGluProSerGlyThrArgHisIleAspAsnThrAlaSerSer 811
QY 3120 ACTCTCTCTCTGCTCCCAACATCGCAGTCAGAGACAGACAGCTTACCTCCATCTCTCCACC 3179
Db 812 SerSerSerCysLeuHisGlnSerAlaValArgLysThrAlaThrAlaHisLeuSerThr 831
QY 3180 TCTAAGAGACAGTCATCTCTCAGCCACGCGAGTGGAA 3215
Db 832 SerLysArgGlnSerSerSerGlyHisAlaValGlu 843
```

RESULT 2

```
Q67907_HPBVO
ID Q67907_HPBVO PRELIMINARY; PRT; 832 AA.
AC Q67907_
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DE 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE DNA polymerase (Fragment).
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sub-type ayw;
RX MEDLINE=94079539; PubMed=8257295;
RA Preisler-Adams S., Schlayer M.J., Peters T., Hettler F., Gerok W.,
RA Rasenack J.
RT "Sequence analysis of hepatitis B virus DNA in immunologically
RT negative infection.";
RL Arch. Virol. 133:385-396(1993).
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; X72702; CAAS1254.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006278; F:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001462; DNAPol_viral_C.
DR InterPro; IPR000201; DNAPol_viral_N.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00336; DNA_pol_viral_C; 1.
DR Pfam; PF00242; DNA_pol_viral_N; 1.
DR Pfam; PF00078; RVT_1; 1.
DR ProDom; PD000814; DNAPol_viral_C; 1.
DR PROSITE; PS50878; RT_POL; 1.
FT NON_TER 1 832
FT NON_TER 832 832
SQ SEQUENCE 832 AA; 93864 MW; EC6872097F04BAAP CRC64;
```

Alignment Scores:

Pred. No.:	8,42e-271	Length:	832
Score:	3745.50	Matches:	726
Percent Similarity:	72.16%	Conservative:	49
Best Local Similarity:	67.60%	Mismatches:	57
Query Match:	63.41%	Indels:	243
DB:	2	Gaps:	2

US-10-761-006a-1 (1-3215) x Q67907_HPBVO (1-832)

```
QY 1 CTCACAACTTCCACCAAGCTCTGTAGATCCAGGCTCAGGGCTATATTTCTCTGC 60
Db 1 LeuHisAsnPheProProAsnSerAlaArgSerGlnGlyGluArgProValPheProCys 20
QY 61 TGTGGGCTCCAGTTCGGACAGTAAACCTGTTCCGACTACTGCTCTCCCATATCGTC 120
Db 21 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerHisIleVal 40
```

QY	121	AATCTTCTCGAGACTGGGGACCTGCGACCGAAATGGAGAACAACTCAGGATTCCT	180
DB	41	AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisIleArgIlePro	60
QY	181	AGGACCCCTGCTCGTGTACAGCGGGTTTTCTCGTTGACAGAACTCCTCACATACC	240
DB	61	ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr	80
QY	241	GCAGAGTCTAGACTCTG-GTGGACTTCTCTCAATTTCTAGGGGAGACCCACGCTGTT	299
DB	81	AlaGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnHisArgVal-Se	100
QY	300	CTGGCCAAATTCGAGTCCCAACCTCAATCACCACCACTCTGTCCTCCAAATTT	359
DB	100	rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe	120
QY	360	GTCTCGCTATCGCTGATGTGTCGGCGGTTTTATCATATTCCTCTCATCCTGCTGC	419
DB	120	userTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaSe	140
QY	420	TATGCTCATCTCTTTGTTGTTCTTCTGCACTACCAAGGTATGTTGCCCGTTTGCTTC	479
DB	140	rMetProHisLeuLeuValGlySerThrGlyLeuSerArgTyrValAlaArgValSerSe	160
QY	480	TACTTCCAGAAACATCAACACAGCACGGGGCCATGCAAGACCTGCACGACTCTGCTC	539
DB	160	rAsnSerArgIlePheAsnHisGlnArgGlyThrMetGlnAsnLeuHisAspTyrCysSe	180
QY	540	AGGAAACTAGCTTCCCTCTGTTGCTGTACAAAACCTTCGGACGGAAACTGCACCTT	599
DB	180	rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrGlnThrPheGlyArgLysLeuHisLe	200
QY	600	GTATTCCTCCATCCTCGGCTTTCGCAAGATTCTATGGGAGTGGGCTCAGTCC	659
DB	200	uTyrSerHisProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr	220
QY	660	GTTCCTCGGCTCAGTTACTAGTCCATTGTTTCAGTGGTTCGTAGGGCTTTCCTCCA	719
DB	220	oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi	240
QY	720	CTGTTTGGCTTCAGTTATATGATGATGTTGTTGGGGCGAAGTCTGTACACATCT	779
DB	240	sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe	260
QY	780	TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCC	839
DB	260	uGluSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr	280
QY	840	TAATAAAACAAACGTTGGGCTTACCTCCCTTAACCTCATGGGATATGTAATTGGAAGTTG	899
DB	280	oAsnLysThrLysArgTrpGlyTyrSerLeuHisPheMetGlyTyrValIleGlySerTy	300
QY	900	GGGTACTTTACCGCAGGAACATATTGTACTAAACCTCAAGCAATGTTTCGAAAACCTGCC	959
DB	300	rGlySerLeuProGlnAspHisIleIleGlnLysIleLysGluCysPheArgLysLeuPr	320
QY	960	TGTAAATAGACCTATTGATTGGAAGTATGTCAAAGAAATTGTGGGTCTTTTGGGCTTGC	1019
DB	320	oIleAsnAsnGProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl	340
QY	1020	TGCCCCCTTTTACAAATATGGCTATCTGCTTGTATGCCCTTTATATGATGTATACAATC	1079
DB	340	aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe	360
QY	1080	TAAAGCAGGCTTTCACCTTCTCGCAACTTACAAGGCTTCTGTGTAAACAATATCTGAA	1139
DB	360	rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs	380
QY	1140	CTTTTACCCCGTTGCCGGCAACGGTCCGTCTCTGCCAAGTGTGTTGCTGACGCAACCCC	1199
DB	380	nLeuTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr	400
QY	1200	CACTGATGGGGCTTTGGCCATAGGCCATCAGGCGATGGGTGGAACCTTTCTGGCTCCTCT	1259
DB	400	oThrGlyTrpGlyLeuValMetMetGlyHisGlnArgMetArgGlyThrPheLeuAlaProIle	420
QY	1260	GCGGATCCATCTACTCGGAACCTCCTAGCAGCTTGTGTTGCTGCAGCCGCTCTGGAGCAA	1319
DB	420	uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaAs	440
QY	1320	ACTTATCGGAACCCGCAACTCTGTTGTCTCTCTCGGAAATACACCTCTCTTTCCATGGCT	1379
DB	440	nIleLeuGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe	460
QY	1380	GCTAGGTGTGCTGCAACTGATCTCTCGGGGAGCGTCTTGTGCTAGTCCCGTCGGC	1439
DB	460	uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl	480
QY	1440	GCTGAATCCCGGAGACCGCTCTCGGGCGGCTTTGGGGCTCTACCGTCCCTCTTCTTCA	1499
DB	480	aleuAsnProAlaAspAspProSerArgLysArgLeuGlyIlePheArgProLeuLeuAr	500
QY	1500	TCTGCCGTTTCGGCCGACACCGGGCGCACCTCTCTTTACGGCGTCTCCCGTATGTGCC	1559
DB	500	gLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaAspSerProSerValPr	520
QY	1560	TTCTCATCTCGCGGACCGTGTGCACTTCGTTTCACTCTGCACTGCGATGGAGACCACC	1619
DB	520	oSerHisLeuProValArgValHisPheAlaSerProLeuHisValAlaTyrArgProPr	540
QY	1620	GTGACGACGCCAGGTCTTGCCCAAGCTCTTATATAGAGACCTCTTGGACTCTCAGCA	1679
DB	540	o-----	540
QY	1680	ATGTCAACGACCGACTTTGAGGCATCTTCAAAGACTGTGTGTTTAAAGACTGGGAGAG	1739
DB	540	-----	540
QY	1740	TTGGGGAGGAGATTAGGTAAAGATTATGTACTAGGAGGCTGTAGCAATAAATTTGGTC	1799
DB	540	-----	540
QY	1800	TGTTCAACGACCACTGCAACTTTTCTCTCTGCTAATCATCTCATGTTCATGTCCTTA	1859
DB	540	-----	540
QY	1860	CTGTTCAAGCCTCCAAGCTGTGCTTGGGTGGCTTTGGGACATGCAATTCACCCGTATA	1919
DB	540	-----	540
QY	1920	AAGAAATTTGGAGCATCTGCTGAGTTACTCTCTTTTTCCTTCTGACTTCTTTCCGTCTA	1979
DB	540	-----	540
QY	1980	TTGAGATCTCTCCGACACCGCTCTGCTGTATCGGAGGCGCTTAGAGTCTCCGGAAC	2039
DB	540	-----	540
QY	2040	ATTGTTGCGCTCACCATACAGCACTCAGCAAGCTATTTTGTGTTGGGTGAGTTGATGA	2099
DB	540	-----	540
QY	2100	ATCTGGCCACTGGGTGGGAAGTAATTTGGAAGATCCAGCATCCAGGGAATTAGTAGTCA	2159
DB	540	-----	540
QY	2160	GCTATGTCAAAGTTAATATGGGCCTAAAACTCAGACAAATATTGTGGTTTTCATTTCT	2219
DB	540	-----	540
QY	2220	GTCTTACTTTTGGGAAGAGAAACGTGTTCTTGAGTACTTGTGTATCTTTTGGAGTGTGATTC	2279
DB	540	-----	540
QY	2280	GCACCTCTACCGCTTTCAGACCACCAAAATGCCCTTATCTTATCAACACTTCCGGAAACTA	2339

```
Db 541 -----Met-----ProLeuSerTyrGlnHisPheArgLysLeu 551
Qy 2340 CTGTTGTTAGACGACGAGGAGGCTCCCTAGAGAAGAACTCCCTCGCTCGCAGACGAA 2399
Db 552 LeuLeuLeuAspGluAlaGlyProLeuGluGluGluLeuProArgLeuAlaAspGlu 571
Qy 2400 GGTCTCAATCGCGCTCGCAGAGATCTCAATCTCGGGAATCTCAAGTTAGTATTCCT 2459
Db 572 AspLeuAsnArgValAlaGluAspLeuAsnLeuGlyAsnLeuAsnValSerIlePro 591
Qy 2460 TGGACTCATAGGTGGGAATCTTACTGGGCTTTATTTCTTCTACTGTACCTGCTTTAAT 2519
Db 592 TrpThrHisLysValGlyAsnPheThrGlyLeuTyrSerSerThrValProValPheAsn 611
Qy 2520 CCGAGTGGCAAAATCTCTCTCTTCTCACAATTCATTACAGAGGACATTATTATAGA 2579
Db 612 ProHisTrpLysThrProSerPheProAsnIleHisLeuHisGlnAspIleLysLys 631
Qy 2580 TGTCAACATATGTGGGCTCTTACAGTTAATGAAAGAGATTAATAATTAATG 2639
Db 632 CysGluGlnPheValGlyProLeuThrValAsnGluLysArgLeuGlnLeuIleMet 651
Qy 2640 CCTGCTAGGTATCTTACCTTACTAATATATTTGCCCTTAGACAAGGCATTAAACCG 2699
Db 652 ProAlaArgPheTyrProAsnValThrLysTyrLeuProLeuAspLysGlyIleLysPro 671
Qy 2700 TATTATCTCAACATGCAATTAATCAATCTTCTCAAACTAGGCAATTTTACATACTCTG 2759
Db 672 TyrTyrProGluHisLeuValAsnHisTyrPheGlnThrArgHisTyrLeuHisThrLeu 691
Qy 2760 TGGAGCTGGCATCTTATATAGAGAGAACAATACACGCGGCTCAATTTTGGTGCTCA 2819
Db 692 TrpLysAlaGlyIleLeuTyrLysArgGluThrThrArgSerAlaSerPheCysGlySer 711
Qy 2820 CCATATCTTGGGAACAGAGCTACACATGGGAGGTGGTCTTCCAACTCGACACAG 2879
Db 712 ProTyrSerTrpGluGlnGluLeuGln----- 720
Qy 2880 CATGGGAGCAATCTGCTGTTCCCAATCTCTGGGATCTTTCCTCCGATCACCAGTTGGA 2939
Db 721 HisGlyAlaGluSerPheHisGlnGlnSerSerGlyIleLeuSerArgProProValGly 740
Qy 2940 CCCTGCTCGGAGCAACTCAACAATCCAGATGGGACTTCAACCCCAACAGGATCA 2999
Db 741 SerSerLeuGlnSerLysHisArgLysSerArgLeuGlyLeuGlnSerGlnGlnHis 760
Qy 3000 CTGCGCAGGCAATCAGGTAGGAGTGGGAGCATTCGGCCAGGTTTCAACCCACCA 3059
Db 761 LeuAlaArgArgGlnGlnGlyArgSerTrpSerIleArgAlaGlyIleHisProThrAla 780
Qy 3060 CGCGGCTCTTTGGGGGGAGCCCTCAGGCTCAGGCAATTTGACAAACAGTCCAGCAGC 3119
Db 781 ArgArgProPheGlyValGluProSerGlySerGlyHisAsnThrAsnLeuAlaSerLys 800
Qy 3120 ACCTCTCTGCTCCCAATCGCAGTCCAGGAGACAGCTACTCCCATCTCTCCACC 3179
Db 801 SerAlaSerCysIleTyrGlnSerProValArgLysAlaIleTyrProAlaValSerThr 820
Qy 3180 TCTAAGAGACAGTCACTCAGCCAGCCAGCGAGTGA 3215
Db 821 PheGluLysHisSerSerSerGlyHisAlaValGlu 832
RESULT 3
Q69590_HPBVO
ID Q69590_HPBVO PRELIMINARY; PRT; 842 AA.
AC Q69590;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polymerase (Fragment).
GN Name=P;
OS Hepatitis B virus.
```

```
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Genotype e;
RX MEDLINE=94120723; PubMed=8291231;
RA Norder H., Courouce A.M., Magnus L.O.;
RT "Complete genomes, phylogenetic relatedness, and structural proteins
RT of six strains of the hepatitis B virus, four of which represent two
RT new genotypes.";
RL Virology 198:489-503 (1994).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Genotype e;
RA Norder H.M.L.;
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; X75657; CAAG3339.1; -: Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001462; DNAPol_viral_N.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00336; DNA_pol_viral_C; 1.
DR Pfam; PF00242; DNA_pol_viral_N; 1.
DR Pfam; PF00078; RVT_1; 1.
DR ProDom; PD000814; DNAPol_viral_C; 1.
DR PROSITE; PS00878; RT_POL; 1.
FT CHAIN <1> >842 Potential.
FT NON_TER 1
FT NON_TER 842
FT SEQUENCE 842 AA; 94812 MW; 45C3CD71CF5A4FE7 CRC64;
SQ
Alignment Scores:
Pred. No.: 6,51e-267 Length: 842
Score: 3693.50 Matches: 719
Percent Similarity: 71.85% Conservative: 52
Best Local Similarity: 67.01% Mismatches: 70
Query Match: 62.53% Indels: 233
DB: 2 Gaps: 2
US-10-761-006A-1 (1-3215) x Q69590_HPBVO (1-842)
Qy 4 CACAACATTCACCAAGCTCTGCTAGATCCAGGCTGAGGGCTATATTTCTCTGCTGG 63
Db 2 HisAsnIleProProSerSerAlaGlySerGlnSerLysArgProValPheSerCysTrp 21
Qy 64 TGCTCCAGTTCGGGAACAGTAAACCTGTTCCGACTACTCTCTCCATATCGTCAAT 123
Db 22 TrpLeuGlnPheArgSerGluProCysSerAspTyrCysLeuThrHisLeuValAsn 41
Qy 124 CTTCTCGAGGACTGGGGACCTCGACCGAATCGGAGAACACATCAGGATTCCTAGG 183
Db 42 LeuLeuGluAspTrpGlyProCysThrGluHisGlyLysHisIleArgIleProArg 61
Qy 184 ACCCTGCTGCTGTATACAGCGGGGTTTTCTGTTGACAGAATCTCTCAATACCGCA 243
Db 62 ThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThrAla 81
Qy 244 GAGTCTAGACTCTG-GTGACTTCTCTCAATTTCTAGGGGAGACCCACGCTGCTCTG 302
Db 82 GluSerArgLeuValValAspPheSerGlnPheSerArgLysSerArgVal-SerTr 101
Qy 303 GCCAAATTCGAGTCCCAACCTCCCAATCACCACCAACCTCTGTCTCCCAATTTGTC 362
Db 101 pProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeuSe 121
Qy 363 CTGCTATACGCTGAGATGTGTCTGGGGGTTTTTATCATATTCCTCTTCATCTGCTGCTAT 422
```


|||||
583 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
QY 900 GGGTACTTTACCGCAGCAATATCTACTAAACCTCAAGCAATGTTTTCGAAACCTGCC 959
Db 603 pGlyThrLeuProGlnGlnHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPr 623
QY 960 TGTAAATAGACCTATTGATTGGAAGATATGTCAAAGAATTGTGGGTCTTTTGGGCTTTC 1019
Db 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
QY 1020 TGGCCCTTTTACACATGCGCTATCTCGCTGTATGCTTATATGATGATGATATCAATC 1079
Db 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
QY 1080 TAAGCAGGCTTTTCACTTTCTCGCAACTTACAGGCTTTCTGTGTAAACAATATCTGAA 1139
Db 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 683
QY 1140 CCTTTACCCCGTTTGGCCGCAACGGTCCGGTCTCTGCCAAGTGTGTTGTGAGCAACCC 1199
Db 683 nLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
QY 1200 CACTGATGGGCTTGGCCATAGGCCCATCAGCGCATGGCTGGAACTTTCTGGCTCTCT 1259
Db 703 oThrGlyTyrGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLe 723
QY 1260 GCGGATCCATCTGCGGAACCTCTAGCAGCTTGTGTTGCTGCGCAGCGCTCTGGAGCAA 1319
Db 723 uProIleHisThrAlaGlnLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys 743
QY 1320 ACTTATCGGAACCGCAAACTCTGTGCTCTCTCGGAATACACCTCTCTTCCATGGCT 1379
Db 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 763
QY 1380 GCTAGGGTGTGTGCGCAACTGATCTGCGCGGACGTCTCTTGTCTACGTCCCGTCGCG 1439
Db 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
QY 1440 GCTGATCCCGGAGCAGCCGCTCTCGGGCGGTTTGGGGCTACCGTCCCTCTTCA 1499
Db 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 803
QY 1500 TCTGCGGTTCCGCGCCACACACCGCGCACCTCTCTTTACGGGTCTCCCGGTATGTGCG 1559
Db 803 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
QY 1560 TTCTCATCTCGCGGACCGGTGACCTTCGCTTACCTCTGCACTGCGCGCATGGAGACCACC 1619
Db 823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaLysArgProPr 843
QY 1620 G 1620
Db 843 o 843

RESULT 5

Q4FD83_HPBVO PRELIMINARY; PRT; 843 AA.
ID Q4FD83_HPBVO PRELIMINARY;
AC Q4FD83;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Polymerase.
GN Name=P;
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HK343;
RX PubMed=15897987; DOI=10.1086/430324;
RA Chan H.L., Tsui S.K., Tee C.H., Ng E.Y., Au T.C., Yuen L.,

RA Bartholomeusz A., Leung K.S., Lee K.H., Locarnini S., Sung J.J.;
RT "Epidemiological and virological characteristics of 2 subgroups of
RT hepatitis B virus genotype C";
RL J. Infect. Dis. 191:2022-2032(2005).
DR EMBL; DQ089793; AAZ05280.1; -; Genomic DNA.
SQ SEQUENCE 843 AA; 94494 MW; BBS0233DFEBA4E29 CRC64;
Alignment Scores:
Pred. No.: 8, 92e-202 Length: 843
Score: 2822.00 Matches: 530
Percent Similarity: 98.5% Conservative: 3
Best Local Similarity: 97.9% Mismatches: 7
Query Match: 47.7% Indels: 2
DB: 2 Gaps: 0

US-10-761-006a-1 (1-3215) x Q4FD83_HPBVO (1-843)

QY 1 CTCACAACATTCCACCAAGCTCTGCTAGATCCCAAGGTGAGGGCCCTATATTTCTCTGC 60
Db 304 LeuHisAsnIleProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCys 323
QY 61 TGGTGGCTCCAGTTCGGAACAGTAAACCTGTTCCGACTACTGCTCTCCATATCGTC 120
Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 343
QY 121 AATCTTCTCGAGGACTGGGGACCTCGCACCAACATGAGAACACAAATCAGGATTCT 180
Db 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
QY 181 AGGACCCCTGCTGCTGTTACAGGGGGGTTTTCTGTTGACAAGAACTCTCACAAATACC 240
Db 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
QY 241 GCAGAGTCTAGACTCTG-GTGGACTTCTCAATTTCTAGGGGAGGAGCACCACGTGTC 299
Db 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
QY 300 CTGGGCCAAAATTCGAGTCCCAACTCCAATCACTCAACCACTCTTGTCTCTCAATTT 359
Db 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
QY 360 GTCTGGCTATCGCTGGATGTCTGCGCGGTTTTATCATATTCCTTCTCATCTCTGCTGC 419
Db 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
QY 420 TATGCTCATCTTCTTGTGGTCTTCTCGGACTACCAAGGTATGTTGGCCGTTCTCTC 479
Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 463
QY 480 TACTTCCAGGAACATCAACACCGAGCGGCCCATGCAAGACCTGCAAGCTCTCTGCTC 539
Db 463 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
QY 540 AAGGAAACTCTACGTTTCCCTCTTGTGCTGTACAAAACCTTTCGGACCGGAAACTGCAT 599
Db 483 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 503
QY 600 GTATTCCCATCCCATCATCTGGGCTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCC 659
Db 503 uTyrSerHisProIleLeuLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
QY 660 GTTTCCTCTGGCTCAGTTTACTAGTGCATTTGTTTGTGTTTGTGTTTGTGTTTGTGTT 719
Db 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 543
QY 720 CTGTTTGGCTTTTCACTTATGATGATGTGTATTTGGGGCGGAAAGTCTGTACAACTCT 779
Db 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
QY 780 TGAGTCCCTTTTACCTCTATATACAAATTTCTTTTGTCTTTTGGGTATACATTTAAACC 839
Db 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583


```
Db      200  utySerHisProIleLeuGlyPheArgylsileProMetGlyValGlyLeuSerPr 220
QY      660  GTTTCCTCCGGCTCAGTTACTAGTCCCAATTTGTTTCAGTGGTTCGTAGGGCTTCCCA 719
Db      220  oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProH 240
QY      720  CTGTTGGCTTTTCAGTTATATGATGATGTTGTTGGGGCGAAGTCTGTACAACATCT 779
Db      240  sCyLeuAlaPheSerTyMetAspAspValValLeuGlyAlaLysSerValGlnHis 260
QY      780  TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACC 839
Db      260  uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAs 280
QY      840  TAATAAAACCAACGTTGGGGCTACTCCCTTAACCTTCATGGGATATGTAATTTGGA 899
Db      280  oAnlyThrLysArgTrpGlyTySerLeuAsnPheMetGlyTyValIleGlySerTr 300
QY      900  GGTACTTTTACCGCAGCAATATTTGTTACTTAAACTCAAGCAATGTTTTCGAAACTGC 959
Db      300  pGlyThrLeuProGlnGluHisIleValGlnLysLeuLysGlnCysPheArgLysLe 320
QY      960  TGTAAATAGACCTATTGATTGGAAGTATGTCMAAGATTGGGCTTTTGGGCTTTCG 1019
Db      320  oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuGlyPheAl 340
QY      1020  TGGCCCTTTTACACAATGGCTATCTGCTTGTATGCTTGTATATGTCATGTATACAAT 1079
Db      340  aAlaProPheThrGlnCysGlyTyProAlaLeuMetProLeuTyAlaCysIleGln 360
QY      1080  TAAGCAGGCTTTTCACTTTCTCGCAACTTCAAGGCTTTCTGTGTAAACAATATCTGA 1139
Db      360  rLysGlnAlaPheThrPheSerProThrTyLysAlaPheLeuCysLysGlnTyLeuAs 380
QY      1140  CCTTTACCCGTTTGGCCGCAACGGTCCGGTCTCTGCCAAGTGTTCGTGAGCAACCC 1199
Db      380  nLeuTyProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThr 400
QY      1200  CACTGATGGGCTTCGGCATAGGCCATCAGCGCATGGCTGGCACTTCTGCTCCTCT 1259
Db      400  oThrGlyTrpGlyLeuAlaIleGlyHisArgMetArgGlyThrPheValAlaProLe 420
QY      1260  GCCGATCCCATCTCGCAACTCTAGCAGCTTCTTTTGTCTCGCAGCGGTCTGGAGCAA 1319
Db      420  uProIleHisThrAlaGluLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLy 440
QY      1320  ACTTATCGGAACCGACACTCTGTTGCTCTCTCGGAATATACCTCCTTTCATGGCT 1379
Db      440  sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyThrSerPheProTrpLe 460
QY      1380  GCTAGGGTGTCTGCCCAACTGGATCTCGCGGAGCTCTCTTGTCTAGTCCCGTCGGC 1439
Db      460  uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyValProSerAl 480
QY      1440  GCTGAATCCCGCGGACGACCGTCTCGGGCCGCTTTGGGGCTCTACCGTCCCTTCTCA 1499
Db      480  aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyArgProLeuLeuH 500
QY      1500  TCTGCGGTTCCGGCGGACGACGCGGCGCACTCTCTTTAGCGGTCTCCCGGTATGTC 1559
Db      500  sLeuProPheArgProThrThrGlyArgThrSerLeuTyAlaValSerProSerValPr 520
QY      1560  TTTCTATCTCCGCGACCGTGTGCACTTCGCTTCACTCTCAGCTCGCATGGAGACCA 1619
Db      520  oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgPro 540
QY      1620  G 1620
Db      540  O 540
RESULT 7
Q9QN49_HPBVO
ID   Q9QN49_HPBVO PRELIMINARY;      PRT;   843 AA.
```

```
AC      Q9QN49;
DT      01-MAY-2000 (TRENBLrel. 13, Created)
DT      01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT      01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE      Polymerase.
OS      Hepatitis B virus.
OC      Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
ON      NCBI_TaxID=10407;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=genotype C;
RX      MEDLINE=20528425; PubMed=11074476;
RA      DOI=10.1002/1096-9071(200012)62:4<471::AID-JMV12>3.3.CO;2-O;
RA      Rokuhara A., Tanaka B., Yagi S., Mizokami M., Hashikura Y.,
RA      Kawasaki S., Kiyosawa K.;
RT      "De novo infection of hepatitis B virus in patients with orthotopic
RT      liver transplantation: analysis by determining complete sequence of
RT      the genome.";
RL      J. Med. Virol. 62:471-478(2000).
DR      EMBL; AB026815; BAA85065.1; -; Genomic_DNA.
DR      GO; GO:0003677; F:DNA binding; IEA.
DR      GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR      GO; GO:0004519; F:endonuclease activity; IEA.
DR      GO; GO:0016787; F:hydrolase activity; IEA.
DR      GO; GO:0004523; F:ribonuclease H activity; IEA.
DR      GO; GO:0003723; F:RNA binding; IEA.
DR      GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR      GO; GO:0016740; F:transferase activity; IEA.
DR      GO; GO:0006278; F:RNA-dependent DNA replication; IEA.
DR      InterPro; IPR001462; DNAPol_viral_C.
DR      InterPro; IPR00201; DNAPol_viral_N.
DR      InterPro; IPR000477; RVTse.
DR      Pfam; PF00336; DNA_pol_viral_C; 1.
DR      Pfam; PF00242; DNA_pol_viral_N; 1.
DR      Pfam; PF00078; RVT_1; 2.
DR      ProDom; PD000814; DNAPol_viral_C; 1.
SQ      SEQUENCE   843 AA;  94462 MW;  9D741241062DEAAE CRC64;
```

Alignment Scores:

Pred. No.:	9,92e-201	Length:	843
Score:	2808.00	Matches:	527
Percent Similarity:	98.52%	Conservative:	6
Best Local Similarity:	97.41%	Mismatches:	7
Query Match:	47.54%	Indels:	2
DB:	2	Gaps:	0

US-10-761-006A-1 (1-3215) x Q9QN49_HPBVO (1-843)

```
QY      1  CTCACCAACATTCACCAAGCTCTGCTAGATCCAGGGTGAGGGGCTATATTTCTCTGC 60
      |||...|||
Db      304  LeuHisHisIleProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCys 323
QY      61  TGTGTGCTCCAGTTCGGGAACAGTAAACCTGTTCCGACTACTGCTCTCCCATATCGTC 120
      |||...|||
Db      324  TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrcysLeuThrHisIleVal 343
QY      121  AATCTTCTCGAGACTGGGGACCTCGACCAATGAGAGAACACAAACATCAGGATTCCT 180
      |||...|||
Db      344  AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
QY      181  AGGACCCCTCGTCTGTTTACAGCGGGGTTTTTCTCGTTTGACAAGAATCTCTCAATACC 240
      |||...|||
Db      364  ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
QY      241  GCAGACTTAGACTCTG-GTGGACTTCTCTCAATTTTCTAGGGGGAGCACCCACGTGTT 299
      |||...|||
Db      384  ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
QY      300  CTGGCCAAAATTCGACAGTCCCAACCTCCAATCACTCACCACTCTTGTCTCTCCCAATTT 359
      |||...|||
Db      403  rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
QY      360  GTCTCGGTATCGCTGGATGTGTCTCGGGGTTTTTATCATATTTCTCTTCTATCTCTGTC 419
```



```
Db 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
Qy 420 TATGCTCATCTCTTGTGGTCTTCTGGAACCAAGGTATGTTGCCGGTTGTCTCTC 479
Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 463
Qy 480 TACTTCAGGAACATCAACACCAAGCAGCGGGCCATGCAAGACTGCAAGCTGCTCTCTC 539
Db 463 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
Qy 540 AAGGAACACTTACGCTTCTCTGTTGCTGTACAAAACCTTCGACGGGAAACTGCACCTT 599
Db 483 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 503
Qy 600 GTATTCCTCATCCATCATCTCGGCTTTCGCAAGATTCCTATGGAGTGGCCCTCAGTCC 659
Db 503 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
Qy 660 GTTCTCTGGCTCAGTTACTAGTGCATTTGTTTCAGTGTTCGTAGGGCTTCCCGCA 719
Db 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi 543
Qy 720 CTGTTTGGCTTTCAGTTATATGATGATGTGATTTGGGGCGAAGTCTGTACAACTCT 779
Db 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
Qy 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTCTCTGTTGGGTATACATTTAAACCC 839
Db 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
Qy 840 TAATAAAACCAAGCTTGGGCTACTCCCTTAACCTTCATGGGATATGATTAATGGAATTG 899
Db 583 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
Qy 900 GGGTACTTTTACCGCAGCAACATATTGTAATAAATCAAGCAATGTTTTCGAAACCTGCC 959
Db 603 pGlyThrLeuProGlnGlnHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPr 623
Qy 960 TGTAAATPAGACTTATTGATGGAAAGTATGTCAAAGAAATTTGGGGTCTTTGGGCTTTC 1019
Db 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
Qy 1020 TGCCCTTTTACCAAGTGGCTATCTGCTGCTGATGCTTATATGATGATGATGATGATGAT 1079
Db 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
Qy 1080 TAAGCAGCTTTTCACTTTCTCGCCAACTTACAGCGCTTTCTGTGTAAACAATATCTGAA 1139
Db 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuHi 683
Qy 1140 CTTTATCCCGTTTCCCGCAACCGTCCGGTCTCTGCAAGTGTGTTGCTGACGCAACCCC 1199
Db 683 sLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
Qy 1200 CACTGGATGGGCTTGGCCATAGGCCATCAGCCATGCGCTGGAACCTTTCTGGCTCTCTCT 1259
Db 703 oThrGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLe 723
Qy 1260 GCCGATCCATCTCGGAACTCTCAGCAGCTTGTGTTGCTCGCAGCGGCTCTCGAGCAAA 1319
Db 723 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLy 743
Qy 1320 ACTTATCGGAACCGCAACTCTGTGTCTCTCTCGGAAATACACCTCTCTTTCATCGCT 1379
Db 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 763
Qy 1380 GCTAGGTGTGCTGCCAACTGGATCTCGCGCGGAGCTCTTGTCTACGTCGCGTCCGCGC 1439
Db 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
Qy 1440 GCTGAATCCCGCGGAGCAACCGTCTCGGGCGGCTTGGGGCTCTACCGTCCCTCTCTCA 1499
```

```
Db 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 803
Qy 1500 TCTGCCGTTCCGCCGCCACACCGCGGCACCTCTCTTTACGGGCTCTCCCGTATGTGCC 1559
Db 803 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
Qy 1560 TTCTCATCTCGCGGACCGTGTGCACCTGCTTCACCTCTGACGCTGCGATGGAGACACCC 1619
Db 823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 843
Qy 1620 G 1620
Db 843 O 843

RESULT 8
DPOL_HPBVR STANDARD; PRT; 843 AA.
ID DPOL_HPBVR AC P03157;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE P protein [includes: DNA-directed DNA polymerase (EC 2.7.7.7); RNA-
DE directed DNA polymerase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN Names:P;
OS Hepatitis B virus (subtype adr).
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=106820;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX MEDLINE=83158919; PubMed=6300776;
RA Ono Y., Onda H., Sasada R., Igarashi K., Sugino Y., Nishioka K.;
RT "The complete nucleotide sequences of the cloned hepatitis B virus
RT DNA; subtype adr and adv.";
RL Nucleic Acids Res. 11:1747-1757(1983).
CC -1- CATALYTIC ACTIVITY: Deoxynucleoside triphosphate + DNA(n) =
CC diphosphate + DNA(n+1).
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonooester.
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL: V00867; -; NOT_ANNOTATED_CDS; Genomic_DNA.
CC PIR: A00704; JDVLYR.
CC InterPro: IPR001462; DNAPol_viral_C.
CC InterPro: IPR00201; DNAPol_viral_N.
CC InterPro: IPR000477; RVTse.
CC Pfam: PF00336; DNA_pol_viral_C; 1.
CC Pfam: PF00242; DNA_pol_viral_N; 1.
CC Pfam: PF00078; RVT_1; 1.
CC ProDom: PD000814; DNAPol_viral_C; 1.
CC PROSITE: PS50878; RT_POL; 1.
CC DNA replication; DNA-binding; DNA-directed DNA polymerase;
CC Endonuclease; Hydrolase; Multifunctional enzyme; Nuclease;
CC Nucleotidyltransferase; RNA-directed DNA polymerase; Transferase.
CC FT DOMAIN 357 600 Reverse transcriptase.
CC SQ SEQUENCE 843 AA; 94400 MW; A6B2D490839C48BB CRC64;

Alignment Scores:
Pred. No.: 1,18e-200 Length: 843
Score: 2807.00 Matches: 528
Percent Similarity: 98.33% Conservative: 3
Best Local Similarity: 97.78% Mismatches: 8
Query Match: 47.52% Indels: 2
DB: 1 Gaps: 0

US-10-761-006A-1 (1-3215) x DPOL_HPBVR (1-843)
```

QY	4	CACAACATCTCCACCAAGCTCTGCTAGATCCAGGGGTGAGGGGCTATATATTTTCTCGTGG	63
Db	305	HiAaNilProProSerSerAlaAArgSerGlnSerGluGlyProIlePheSerCysTrp	324
QY	64	TGCTCCAGTTCGGGAACAGTAACCTGTTCGAGCTACTGCTCTCCCATATCGTCAAT	123
Db	325	TrpLeuGlnPheArgAenSerLysProCysSerAspTyrCysLeuThrHisIleValAa	344
QY	124	CTTCTCGAGACTGGGAGACCCTGCACCGAATCATGGAGAACACAACTACGAGTTCCTAGG	183
Db	345	LeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisnIleArgIleProArg	364
QY	184	ACCCCTGCTCGTTTACAGCGGGGGTTTTCTCGTTTGACAAGAATCTCTCAATACCGCA	243
Db	365	ThrProAlaArgValThrGlyGlyValPheLeuValAAspLysAenProHisAenThrThr	384
QY	244	GAGTCTAGACTCTG-GTGGACTTCTCTCAATTTTCTAGGGGGAGCACCCACGTGTTCCTG	302
Db	385	GluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-SerTr	404
QY	303	GCCAAATTCGCAGTCCCCCACTCCAACTCACTCACCAACTCTGTGTCTCCCAATTTGTC	362
Db	404	pProLysPheAlaValProAenLeuGlnSerLeuThrAaenLeuLeuSerSerAaenLeuSe	424
QY	363	CTGGCTATCGCTGGATGTCTCGGGGGTTTTTATCATATTTCTCTTCATCTCTGCTCAT	422
Db	424	rTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAaMe	444
QY	423	GCCTCATCTTCTGTGTGTTCTTCTGACTACCAAGGTATGTGGCGGTGTGCTCTCTAC	482
Db	444	tProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSerTh	464
QY	483	TTCCAGAAACATCAACACACAGCAGCGGGCCATGCAAGACTCTGCAGACACTCTGCTCAAG	542
Db	464	rSerArgAaNilAaenHisGlnHisGlyAlaMetGlnAspLeuHisAspSerCysSerAr	484
QY	543	GAAACTCTACGTTTCCCTCTTGTGTGTATCAAAACCTTCGAGCGGAAACTGCACATTGTA	602
Db	484	gAenLeuTyr-ValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLeuTy	504
QY	603	TTCCCATCCCATCATCTCGGCTTTCACAGATTCCTATGGAGTGGGCTCAGTCGGTT	662
Db	504	rSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerProPh	524
QY	663	TCTCTCGCTCAGTTTACTAGTGCATTGTTCAGTGTGTTCTGATGGGCTTTCGCCCACTG	722
Db	524	eLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHisCys	544
QY	723	TTTGGCTTTCAGTTATATGATGATGTGTATTTGGGGCGGAAGTCTGTACAACTTTTGA	782
Db	544	sLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLeuGl	564
QY	783	GTCCCTTTTACCTTATATACAAATTTTCTTTGTCTTTGGGTATACATTTAAACCCATA	842
Db	564	uSerLeuPheThrSerIleThrAenPheLeuLeuSerLeuGlyIleHisLeuAaenProAs	584
QY	843	TAAACCAAGTGTGGGCTACTCCCTTAACCTTCATCGGATATGCTATTCGAGTGGGG	902
Db	584	nLysThrLysArgTrpGlyTyrSerLeuAaenPheMetGlyTyrValIleGlySerTrpGl	604
QY	903	TACTTTTACCGCAGAAACATATTGTACTTAAAACTCAAGCAATGTTTTTCGAAAACTGCCTG	962
Db	604	yThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuProVa	624
QY	963	AAATAGACCTTATGATTGGAAAGTATGTCAAGAATTTGTGGGTCTTTTGGGCTTTGCTGC	1022
Db	624	lAaenSerProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAlaAl	644
QY	1023	CCCTTTTACAAATGCGCTATCTGCTTGATCGCTTATATGTCATGATATACAACTCAA	1082
Db	644	aProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSerLys	664
QY	1083	GCAGGCTTTCATCTTCTCGCCAACTTACAAAGGCCCTTTCTGTGTGTAAACAAATCTGAACCT	1142

Db	664	sGlnAlaPheThrPheserProthrTyrLySaLaPheLeuCysLyeGlnTyrLeuaenLe	684
Qy	1143	TTACCCCGTGTCGCCGGCAACGGTCCGGTCTCTGCCAAGTGTGTTGCTGCAGCAACCCCCCAC	1202
Db	684	utyProVaLaLaargLnrgSerGlyLeuCysGlnValPheAlaAspAlaThrProTh	704
Qy	1203	TGGATGGGGCTTGGCCATAAGCCATCAGCGCATGGCTGGGAACCTTTCTTGCTCTCTCTGCC	1262
Db	704	rGlyTrpGlyLeuAlailegLYHisArgMetArgGlyThrPheValAlaProLeuPr	724
Qy	1263	GATCCATACTGCGGAACCTCTACCAGCTTGTTTTGCTCGCAGCGGGTCTGGAGCAAACT	1322
Db	724	oIleHsThrAlagLeuAlalaCaYsPheAlaArgSerArgSerglyAlaYsLe	744
Qy	1323	TATCGGAACCGACAACCTGTGTTCTCTCTCGGAATAACACCTCCTTTTCCATGGCTGCT	1382
Db	744	uIlleGlyThrAspAnsSerValVallueSrArgLYstYrThrSerPheProTrpLeule	764
Qy	1383	AGGGTGTGCTGCCAACCTGGATCTCGCGCGGACGTCTCTTTGTCTAGCTCCCGTCGGCGCT	1442
Db	764	uGlyCySaLaAlaAsnTrpileLeuargGlyThrSerPheValTyValProSerAlaLe	784
Qy	1443	GAATCCCGGGAGACACCGTCTCGGGGCCGTTTGGGGCTTACCGTCCCCTTTCTCATCT	1502
Db	784	uAnPrCaLaAspProSerArgGlyArGLeuGlyLeuTyArgProLeuLeuLe	804
Qy	1503	GCGGTTCGGCGCACCAACGGGGCGCACTCTCTTAGCGGGTCTCCCGGTATGTCCTTC	1562
Db	804	uProPheArGproThrThrglyargThrSerLeuTyArAlavaSerProSerValProSe	824
Qy	1563	TCATCTGCCGGACCGTGTGCACCTTCGCTTCACCTCTGCAGTCGCATGGAGACACCG	1620
Db	824	rHIsLeuProAspargValHisPheAlaSerProLeuHisValAlaTrpArgProPro	843
RESULT 9			
Q97975 HPBV0 PRELIMINARY; PRT; 540 AA.			
ID	Q97975		
AC	Q97975		
DT	01-FEB-1997	(T-EMBLrel. 02, Created)	
DT	01-FEB-1997	(T-EMBLrel. 25, Last sequence update)	
DT	01-OCT-2003	(TrEMBLrel. 25, Last annotation update)	
OS	ORF [fragment]		
DB	Hepatitis B virus.		
OC	Viruses; Retro-transcribing viruses; Hepadnaviridae;		
OC	Orthohepadnavirus.		
OX	NCBI_Taxid=10407;		
RN	[1]		
NUCLEOTIDE SEQUENCE.			
RR	MEDLINE=95294549; PubMed=7775946;		
RA	Uchida T., Gotoh K., Shikata T.		
RT	*Complete nucleotide sequences and the characteristics of two		
RT	hepatitis B virus mutants causing serologically negative acute or		
RT	chronic hepatitis B.;		
J.	Med. Virol. 45:247-252(1995).		
CC	-1- STMILARY. Contains 1 reverse transcriptase domain.		
DR	EMBL; D16666; BAA04073.1; -; Genomic_DNA.		
DR	GO; GO:0003677; F:DNA binding; IEA.		
DR	GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.		
DR	GO; GO:0004523; F:ribonuclease H activity; IEA.		
DR	GO; GO:0003723; F:rRNA binding; IEA.		
DR	GO; GO:0003964; F:rRNA-directed DNA polymerase activity; IEA.		
DR	GO; GO:0006278; P:rRNA-dependent DNA replication; IEA.		
DR	IncerPro; IPR001462; DNapol_viral_C.		
DR	InterPro; IPR000477; RVTse.		
DR	Pfam; PF00336; DNA_pol_viral_C.1.		
DR	Pfam; PF00078; RVT_1.1.		
DR	ProDom; PD000814; DNapol_viral_C.1.		
DR	PROSITE; PS50878; RT_POL.1.		
FT	NON_TER		
SQ	SEQUENCE	540 AA; 60346 MW; 7BBE285359DD39C5 CRC64;	
Alignment Scores:			

Pred. No.: 1.2e-200 Length: 540
 Score: 2806.00 Matches: 526
 Percent Similarity: 98.34% Conservative: 6
 Best Local Similarity: 97.23% Mismatches: 8
 Query Match: 47.50% Indels: 2
 DB: 2 Gaps: 0

US-10-761-006A-1 (1-3215) x Q97975_HPBVO (1-540)

QY	1	CTCCACAACTTCCACCAAGCTCTGCTAGATCCAGGGTCAGGGCCCTATATTTCTCTGC	60
DB	1	LeuHisanlleProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCys	20
QY	61	TGTGGCTCCAGTTCGGAACAGTAACCTGTTCCGACTACTGCTCTCCCATATGCTC	120
DB	21	TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal	40
QY	121	AATCTTCTCAGGACTGGGACCTCGACCAAGCAATGAGAACACACATCAGGATTCCT	180
DB	41	AsnLeuLeuGlnAspTrpGlyProCysThrGluHisGlyGluHisanlleArgIlePro	60
QY	181	AGGACCCCTCTCTGTTTACAGCGGGGTTTTCTCGTTGACAGAATCCTCAATAATCC	240
DB	61	ArgThrProalaargValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr	80
QY	241	GCAGACTAGACTCTG-GTGGACTTCTCAATTTTCTAGGGGAGACCCACGTTGTC	299
DB	81	ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se	100
QY	300	CTGGCCAAATTCGAGTCCCACTCCCACTCACTCACTCACTCACTCTGTCCTCAATTT	359
DB	100	rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe	120
QY	360	GTCCTGCTATCGTGTGATGCTGGGGGTTTTATCATATCTCTTCTCATCTGCTGC	419
DB	120	uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl	140
QY	420	TATGCTCATCTTCTGTTGTTCTTCTGCACTACCAAGTATGTTGCCGTTGCTCTC	479
DB	140	ametProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe	160
QY	480	TACTTCAGGAACATCAACCAAGCAGCGGGCCATGCAAGACTGCACTGCACTCTGCTC	539
DB	160	rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe	180
QY	540	AAGGAACTCTACGTTTCCCTCTGTTGCTGTGTACAAACCTTCGAGCGGAACTGCACCT	599
DB	180	rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe	200
QY	600	GTATTCCTCCATCCATCTCTGGGCTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCC	659
DB	200	uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr	220
QY	660	GTTTCTCTGGCTCAGTTTACTAGTGCATTTGTTTCACTGTTGTTGAGGCTTCCGCCA	719
DB	220	oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi	240
QY	720	CTGTTTGGCTTTCAGTTTATATGATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTG	779
DB	240	sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe	260
QY	780	TGAGTCCCTTTTACCTATTACCAATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	839
DB	260	uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr	280
QY	840	TAATAAACCAAGCTTGGGCTACTCCCTTAACCTTATGCGATATGATATGGAAGTTG	899
DB	280	oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr	300
QY	900	GGGTACTTTTACCGCAGGAACATATGTTACTTAAACTCAAGCAATGTTTTCGAAACCTGC	959
DB	300	pGlyThrLeuProGlnGluHisIleValLeuLeuLeuLysGlnCysPheArgLysLeuPr	320

QY	960	TGTAATATAGACCTATTGATTGGAAAGTATGTCAAAGAATTGTGGTCTTTTGGGCTTTC	1019
DB	320	oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl	340
QY	1020	TGCCCTTTTACACAAATGCGGTATCTGCTCTTGATGCTTATATGTCATGTATACAATC	1079
DB	340	alaProPheThrGlnCysGlyTyrProalaLeuMetProLeuTyrAlaCysIleGlnSe	360
QY	1080	TAAGCAGGCTTTCACCTTCTCGCAACTTCAAGGCTTCTCTGTGTAAACAATATCTGAA	1139
DB	360	zLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysGlnGlnTyrLeuHi	380
QY	1140	CTTTTACCGGTTGCGGCAAGGTCGCTCTCTGCAAGTGTGTTGCTGACGCAACCCC	1199
DB	380	sLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr	400
QY	1200	CACGTGATGGGCTTGGCCATAGGCCATCAGCCCATGCGTGGAACTTTCTGGCTCTCT	1259
DB	400	oThrGlyTTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLe	420
QY	1260	GCGATCCATACCTGCGAACTCTCTAGCAGCTTCTTGTCTCGCAGCGGTCTGGAGCAA	1319
DB	420	uProIleHisThrAlaGluLeuAlaIleCysPheAlaArgSerArgSerGlyAlaLy	440
QY	1320	ACTTATCGGAACGCACTCTGTTCTCTCTCTCGAAATACACCTCTTCCATGGCT	1379
DB	440	sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe	460
QY	1380	GCTAGGCTGTGCTGCGCAACTGATCTGCGCGGAGCTCTTGTCTTACGCTCCGCTCGGC	1439
DB	460	uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl	480
QY	1440	GCTGAATCCCGCGGACGACCCGCTCTCGGGCCGTTTGGGGCTCTACCGTCCCTTCTTCA	1499
DB	480	aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi	500
QY	1500	TCTGCGGTTCCGCGGACGACGCGGCGGACCTCTCTTACGCGGTCTCCCGGTATGTC	1559
DB	500	sLeuProPheArgProThrGlyArgThrSerLeuTyrAlaValSerProSerValPr	520
QY	1560	TTCTCATCTGCGGACCGGTGTGCACTTCCGCTTCCACCTCTGCACTGCGATGGAGACACC	1619
DB	520	oSerHisLeuProSerArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr	540
QY	1620	G 1620	
DB	540	O 540	

RESULT 10
 Q5KR39_HPBVO PRELIMINARY; PRT; 843 AA.
 ID Q5KR39_HPBVO PRELIMINARY; PRT; 843 AA.
 AC Q5KR39;
 DT 10-MAY-2005 (TrEMBLrel. 30, Created)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
 DE Polymerase.
 GN Names=P;
 OS Hepatitis B virus.
 OS Viruses; Retro-transcribing viruses; Hepadnaviridae;
 OC Orthohepadnavirus.
 ON NCBI_TaxID=10407;
 RN NUCLEOTIDE SEQUENCE.
 RP Sun X., Rokuhara A., Tanaka E., Gad A., Mutou H., Matsumoto A.,
 RA Yoshizawa K., Kiyosawa K.;
 RT "Nucleotide Mutations Associated With Hepatitis B e Antigen
 Negativity.";
 RL J. Med. Virol. 76:170-175 (2005).
 RN NUCLEOTIDE SEQUENCE.
 RP Rokuhara A.;
 RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB198077; BAD86600.1; -; Genomic DNA.

DR GO: GO:0003677; F:DNA binding; IEA.
 DR GO: GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
 DR GO: GO:0004519; F:endonuclease activity; IEA.
 DR GO: GO:0016787; F:hydrolase activity; IEA.
 DR GO: GO:0004523; F:ribonuclease H activity; IEA.
 DR GO: GO:0003723; F:RNA binding; IEA.
 DR GO: GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
 DR GO: GO:0016740; F:transferase activity; IEA.
 DR GO: GO:0006278; P:RNA-dependent DNA replication; IEA.
 DR InterPro: IPR001462; DNAPol_viral_C.
 DR InterPro: IPR000201; DNAPol_viral_N.
 DR InterPro: IPR000477; RVase.
 DR Pfam: PF00336; DNA_pol_viral_C; 1.
 DR Pfam: PF00242; DNA_pol_viral_N; 1.
 DR Pfam: PF00078; RVT 1; 1.
 DR ProDom: PD000814; DNAPol_viral_C; 1.
 SQ SEQUENCE 843 AA; 94446 MW; 80CD75F48581AC2F CRC64;

Alignment Scores:

Pred. No.: 1.4e-200 Length: 843
 Score: 2806.00 Matches: 526
 Percent Similarity: 98.15% Conservative: 5
 Best Local Similarity: 97.23% Mismatches: 9
 Query Match: 47.50% Indels: 2
 DB: Gaps: 0

US-10-761-006A-1 (1-3215) x Q5KR39_HPBVO (1-843)

QY 1 CTCACAACTTCCACCAAGCTCTGCTAGATCCAGGGTGAGGGCGCTATATTTTCCTGC 60
 DB 304 LeuHisAsnIleProProSerSerAlaArgProGlnSerGluGlyProIleLeuSerCys 323
 QY 61 TGTGGCTCCAGTTCGGGAACAGTAACCTGTTCGGAAGTCTCCCATATCGTC 120
 DB 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 343
 QY 121 AATCTTCTCAGAGACTGGGGACCTGACCGAACATATGAGAGAACACAAATCAGGATCTCT 180
 DB 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
 QY 181 AGGACCCCTGCTGCTGTACAGCGGGTTTTCTCGTGTGACAAAGATCCTCACATACC 240
 DB 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
 QY 241 CGAGAGTCTAGACTCTG-GTGGACTTCTCTCAATTTCTAGGGGAGCACCCACGTGTC 299
 DB 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
 QY 300 CTGGCCAAAATTGCGAGTCCCGAACCTCAATCAGTCTCACCAACCTTGTCTCTCCAAATT 359
 DB 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
 QY 360 GTCCGCTCATCTCTGTGGTGTCTGCGGGTTTTATCATATTTCTCTTCATCCCTGCTGC 419
 DB 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
 QY 420 TATGCTCATCTCTTGTGGTGTCTCTGGAAGTATGTTGCGCGTGTGCTTCATCCCTGCTC 479
 DB 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 463
 QY 480 TACTTCCAGAACATCAACACAGACCGGGCCATGCAAGACCTGACAGTCTCTGCTC 539
 DB 463 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
 QY 540 AAGGAACCTAGTGTCTCTGTTGCTGTGTACAAAACCTTCGGACGGAACTGCACATT 599
 DB 483 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisIle 503
 QY 600 GTATTCCTCCATCATCTCTGGGCTTTCGCAAGATTCTATGGAGTGGGGCTCAGTCC 659
 DB 503 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
 QY 660 GTTCTCTCGGCTCAGTTTACTAGTCCAAATTTGTTTCAGTGGTTCGTAGGGCTTTTCCCCCA 719

DB 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 543
 QY 720 CTGTTTGGCTTTTCAGTTATATGATGATGTGTATTTGGGGCGAAGTCTGTACAACATCT 779
 DB 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisIle 563
 QY 780 TGAGTCCCTTTTACCTCTATTTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCC 839
 DB 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
 QY 840 TAATAAAACCAACGTTGGGGCTACTCCCTTAACCTTCATGGGATATGAATTTGGAGTTG 899
 DB 583 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
 QY 900 GGGTACTTTTACCGCAGGAACATATTTGTAATAAATCAAGCAATGTTTTCGAAAACTGCC 959
 DB 603 pGlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPr 623
 QY 960 TGTAAATAGACCTATTGATTGGAAAGTATGTCAAGAATTGTGGGTCTTTTGGGCTTTCG 1019
 DB 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
 QY 1020 TGCCCTTTTACACAAATGTGGCTATCTGCTCCCTTGATGCTTTTATATGATGATATCAATC 1079
 DB 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
 QY 1080 TAAGCAGGCTTTCACCTTTCTCGCCAACTTACAAAGCCTTCTGTGTAAACAATATCTGAA 1139
 DB 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 683
 QY 1140 CCTTTACCCGTTGCCGCAACGTTCCGGTCTCTGCCAAGTGTGTGTGTCGACCAACCC 1199
 DB 683 nLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
 QY 1200 CACTGGATGGGCTTGGCCATAGGCATCAGCGCATGGCTGCAACCTTCTCGCTCCTCT 1259
 DB 703 oThrGlyTrpGlyLeuAlaValGlyHisArgArgMetArgGlyThrPheValAlaProLe 723
 QY 1260 GCGGATCCATCTAGCGGAACCTCTAGCAGCTGTGTTTGTCTCGACCGGCTCTGGAGCAA 1319
 DB 723 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaL 743
 QY 1320 ACTTATCGGAACCGCAACTCTGTGTCTCTCTCGGAATACACTCTCTTCCATGGCT 1379
 DB 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 763
 QY 1380 GCTAGGTGTGTGCGCAACTGATCTGCGCGGAGCTCTTTGTGTAGCTCCGCTCGCGC 1439
 DB 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
 QY 1440 GCTGAATCCCGCGGACGACCCGCTCTCGGGCGGCTTTGGGGCTCTACCGTCCCTTCTTCA 1499
 DB 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 803
 QY 1500 TCTCGGCTTCCGGCCACACCGGGCGCACCTCTCTTTACGCGGTCTCCCGGTATGTGCC 1559
 DB 803 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
 QY 1560 TTCTCATCTGCGGACCGGTGTGCACTTCGCTTCACCTCTGCGACGTGGATGGAGACCAC 1619
 DB 823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 843
 QY 1620 G 1620
 DB 843 o 843

RESULT 11

Q7TDR3 HPBV0
 ID Q7TDR3 HPBV0 PRELIMINARY; PRT; 843 AA.
 AC Q7TDR3-
 DT 01-OCT-2003 (TREMBLrel. 25, Created)
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)

[illegible]


```

Db      763 uLeuGlyCyAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrrValProSerAl 783
Qy      1440 GCTGAATCCCGCGGACACCGCTCTCGGGGCCGTTTGGGGCTCTACCGTCCCTCTTCTCA 1499
Db      783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrrArgProLeuLeuHi 803
Qy      1500 TCTGCCGTTCCGGCCGACACGCGGCGCACCTCTCTTTAGCGGTCTCCCGGTATGGCC 1559
Db      803 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrrAlaValSerProSerValPr 823
Qy      1560 TTCTCATCTCGCGGACCGTGGCACTTCTGCTTCACTCTGCTGCTGCTGCTGCTGCTGCTGCT 1619
Db      823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 843
Qy      1620 G 1620
Db      843 o 843

RESULT 13
Q4FD65_HPBVO PRELIMINARY; PRT; 843 AA.
AC Q4FD65;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Polymerase.
GN Name=P;
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RC NUCLEOTIDE SEQUENCE.
RA CHAN H.L., Tsui S.K., Tse C.H., Ng E.Y., Au T.C., Yuen L.,
RA Bartholomeusz A., Leung K.S., Lee K.H., Locarnini S., Sung J.J.;
RT "Epidemiological and virological characteristics of 2 subgroups of
RT hepatitis B virus genotype C.";
RL J. Infect. Dis. 191:2022-2032 (2005).
DR EMBL; DQ089798; AA205298.1; -; Genomic DNA.
SQ SEQUENCE 843 AA; 94318 MW; 0C749903F281504A CRC64;

Alignment Scores:
Pred. No.: 2.35e-200 Length: 843
Score: 2803.00 Matches: 525
Percent Similarity: 98.34% Conservative: 7
Best Local Similarity: 97.04% Mismatches: 8
Query Match: 47.45% Indels: 2
DB: 2 Gaps: 0

US-10-761-006A-1 (1-3215) x Q4FD65_HPBVO (1-843)
Qy      1 CTCACAAACATCCACCAAGCTCTGTAGATCCAGGGTGAAGGCGCTATATTTCTCTGC 60
Db      304 LeuHisAsnIleProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCys 323
Qy      61 TGGTGGCTCCAGTTCGCGACAGTAACCCGTTCGCACTCTCCCATATCGTC 120
Db      324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrrCysLeuThrHisIleVal 343
Qy      121 AATCTTCTCGAGGACTGGGACCTTGACCCGAAATCGAGAACACACATCAGGATTCCT 180
Db      344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
Qy      181 AGGACCCCTGCTGTTTACAGGGGGGTTTTCTCGTTGACAGAAATCCTCAAAATACC 240
Db      364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
Qy      241 GCAGAGTCTAGACTCTG-GTGGACTTCTCAATTTCTAGGGGGAGCACCCAGTCTTC 299
Db      384 ThrGluSerArgLeuValAlaAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403

```

```

Qy      300 CTGGCCAAAATTTCGAGTCCCGAACCTCAATCACCACCTCTTGTCTCTCCAAATTT 359
Db      403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
Qy      360 GTCTCTGGCTATCTGCTGGATGTGTCTCGCGGTTTTATCATATTCCTCTTCTCATCTCTGCTGC 419
Db      423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrrHisIleProLeuHisProAlaAl 443
Qy      420 TATGCTCATCTTCTTGTGTGTCTTCTGAGCTTACCAAGGTATGTTCGCCGTTTGTCTCTC 479
Db      443 aMetProHisLeuLeuValGlySerSerGlyLeuProGlyTyrrValAlaArgLeuSerSe 463
Qy      480 TACTTCCAGGAACATCAACACCCAGCAGCGGGCCATGCAAGACTGCAAGCTGCAAGCTCTGCTC 539
Db      463 rThrSerArgAsnIleAsnTrpGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
Qy      540 AAGGAAACTCTACGTTTCCCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 599
Db      483 rArgAsnLeuTyrrValSerLeuLeuLeuLeuLeuTyrrLysThrPheGlyArgLysLeuHisLe 503
Qy      600 GTATTTCCCATCCCATCATCTCTGGGCTTTCGCAAGATTCTTATGGGAGTGGGCTCTAGTCC 659
Db      503 uTyrrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
Qy      660 GTTCTCTCTGGCTCAGTTTACTAGTGCATTTGTTTTCAGTGGTTCGTAGGCTTCTCCCTCA 719
Db      523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 543
Qy      720 CTGTTTGGCTTTCAGTTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 779
Db      543 sCysLeuAlaPheSerTyrrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
Qy      780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 839
Db      563 uGluSerLeuTyrrThrAlaIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
Qy      840 TATAAAACAAACGTTGGGCTCTACCTTAACTTCATCGGATATGATTAATTTGGAAGTTG 899
Db      583 oAsnLysThrLysArgTrpGlyTyrrSerLeuAsnPheMetGlyTyrrValIleGlySerTr 603
Qy      900 GGTACTTTTACCCGACGACATATGTTACTTAAACTCAAGCAATGTTTTCGAAACACTGCC 959
Db      603 pGlyThrLeuProGlnGluHisIleValLeuLysIleLysGlnCysPheArgLysLeuPr 623
Qy      960 TGTAAATAGACCTATTGATTGGAAGATATGCTCAAGAAATTTGTGGGTCTTTTGGCTTTTGC 1019
Db      623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
Qy      1020 TGCCCTTTTACAAATGTCGCTATCTGCTGCTGATGCTCTTTATATGATGATGATGATGATGATGAT 1079
Db      643 aAlaProPheThrGlnCysGlyTyrrProAlaLeuMetProLeuTyrrAlaCysIleGlnSe 663
Qy      1080 TTAGCAGGCTTTCACCTTCTCGCCCAACTTACAGGCTTCTCTGTGTAAACAATATCTGAA 1139
Db      663 rLysGlnAlaPheThrPheSerProThrTyrrLysAlaPheLeuCysLysGlnTyrrLeuAs 683
Qy      1140 CCTTTACCCGTTGCGCGCAACGCTCGCTCTCTGCAAGTGTGTTCTGACGCAACCCCC 1199
Db      683 nLeuTyrrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
Qy      1200 CACTGGATGGGCTTGGCCATAGCCCATAGCCGATGCTGGAACTTTTCTGGCTCTCTCTCT 1259
Db      703 oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValAlaProLe 723
Qy      1260 GCCGATCCATCTGCGGAACCTCTAGCAGCTTCTTTTGTCTGCGAGCGGCTCTGAGAGAAA 1319
Db      723 uProIleHisThrAlaGlnLeuLeuAlaCysPheAlaArgSerArgSerGlyAlaLys 743
Qy      1320 ACTTATCGGAACCGACAACTCTGTGTGTCTCTCTCGGAAATATACACCTCTCTTCTCATGGCT 1379
Db      743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrrThrSerPheProTrpLe 763

```

QY 1380 GCTAGGTTGCTGCTCAACTGATCCTCGCGGAGCGTCTTTGTCTACGTCCCGTCGCG 1439
 Db 763 uLeuGlyCyAlaAlaAenTrrPleLeuArgGlyThrSerPheValTyrValProSerAl 783
 QY 1440 GCTGAATCCCGGAGCGCGCTCTCGGGCGGTGGGGCTCTACCGTCCCGCTTCTTCA 1499
 Db 783 aLeuAenProAlaAenAseProSerArgGlyArgLeuGlyLeuPheArgProLeuLeuHi 803
 QY 1500 TCTGCGGTTCCCGCGCGACACCGGGCGCACCTCTCTTTACGCGGTCTCCCGTATGTGCC 1559
 Db 803 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
 QY 1560 TTCTCATCTGCGGACCGGTGTGCACTTCGCTTCACTCTGCACTGCGAGTGAAGACCCACC 1619
 Db 823 oSerHisLeuProAseArgValHisPheAlaSerProLeuHisValAlaTrrArgProPr 843
 QY 1620 G 1620
 Db 843 o 843
 RESULT 14
 Q81165 HPBV0 PRELIMINARY; PRT; 838 AA.
 AC Q81165;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE DNA polymerase.
 OS Hepatitis B virus.
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 OX NCBI_TaxId=10407;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=subtype adr;
 RX MEDLINE=93096607; PubMed=1461746;
 RA Mukaida M.;
 RT "The complete nucleotide sequence of hepatitis B virus, subtype adr
 (SRADR) and phylogenetic analysis.";
 RL Nucleic Acids Res. 20:6105-6105(1992).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=subtype adr;
 RA Uchida T., Aye T.T., Shimojima M., Gotoh K., Shikata T.;
 RT "Full-length nucleotide sequence of a hepatitis B virus (HBV) mutant
 isolated from a patient with acute hepatitis who did not exhibit
 serological markers for HBV infection.";
 RL Int. Hepatol. Commun. 2:70-73(1994).
 DR EMBL; D16685; BAA04072.1; -; Genomic_DNA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0003987; F:DNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0004519; F:endonuclease activity; IEA.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR GO; GO:0004523; F:ribonuclease H activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006278; F:RNA-dependent DNA replication; IEA.
 DR InterPro: IPR001462; DNAPol_viral_C.
 DR InterPro: IPR000201; DNAPol_viral_N.
 DR InterPro: IPR000477; RVase.
 DR Pfam: PF00336; DNA_pol_viral_C; 1.
 DR Pfam: PF00242; DNA_pol_viral_N; 1.
 DR Pfam: PF00078; RV1; 2.
 DR ProDom: PD000814; DNAPol_viral_C; 1.
 SQ SEQUENCE 838 AA, 93902 MW, A686F42940C3B806 CRC64;

Alignment Scores:

Pred. No.: 2,78e-200 Length: 838
 Score: 2802.00 Matches: 525
 Percent Similarity: 98.33% Conservative: 6
 Best Local Similarity: 97.22% Mismatches: 8
 Query Match: 47.44% Indels: 2
 DB: 2 Gaps: 0

US-10-761-006A-1 (1-3215) x Q81165_HPVB0 (1-838)
 QY 4 CACAACTTCCACCAAGCTCTGTAGATCCAGGGTGAGGGCCCTATATTTTCTGCTGG 63
 Db 300 HisAnIlePleProProSerSerAlaAargSerGlnSerGluGlyProIlePheSerCysTrp 319
 QY 64 TGGCTCCAGTTCCGGAAACAGTAAACCTGTTCCGACTACTGCTCTCCCATATCGTCAAT 123
 Db 320 TrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleValAen 339
 QY 124 CTTCTCGAGGACTGGGGACCTGACCCGAAACATGAGAACACAAATCAGGATTCCTAGG 183
 Db 340 LeuLeuGluAseTrpGlyProCysThrGluHisGlyGluHisAenIleArgIleProArg 359
 QY 184 ACCCTGCTCGTGTGTACAGGGGGGTTTTCTCGTTGACAAAGTCTCACAATACCGCA 243
 Db 360 ThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAnThrThr 379
 QY 244 GAGTCTAGACTCTG-GTGGACTTCTCTCAATTTCTAGGGGAGCACCCACGTTCTCTG 302
 Db 380 GluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-SerTr 399
 QY 303 GCCAAAATTCGAGTCCCGACCTCCAACTCACTACCAACCTCTTGTCTCCTCAATTTGTC 362
 Db 399 pProLysPheAlaValProAenLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeuSe 419
 QY 363 CTGGCTATCGTGGATGTCTCGCGCGTTTTATCATATTCCTCTTCTCATCTGCTGCTAT 422
 Db 419 rTrpLeuSerLeuAsePvalSerAlaAlaPheTyrHisIleProLeuHisProAlaAlaMe 439
 QY 423 GCCTCATCTCTTGTGTGTTCTTCTGGACTACCAAGGTATGTCGCGTTTCTCTCTAC 482
 Db 439 tProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSerTh 459
 QY 483 TTCCAGGAAACATCAACACACGAGCGGCCCATGCAAGCTTCGACGACTCTCTGCTCAAG 542
 Db 459 rSerArgAsnIleAenTyrGlnHisGlyThrMetGlnAspLeuHisAseSerCysSerAr 479
 QY 543 GAAACTCTACGTTTCCCTCTCTGCTGTACAAAACCTTCGACGCGAAATCGACATTGTA 602
 Db 479 gAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLeuTy 499
 QY 603 TTCCCATCCCATCATCTCGGGCTTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCCGTT 662
 Db 499 rSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerProPh 519
 QY 663 TCTCTGCTCAGTTTACTAGTGCCTATTTGTTTCAGTGGTTCGTAGGGCTTTCCTCCCATG 722
 Db 519 eLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHisCy 539
 QY 723 TTTGGCTTTTCAGTTATATGATGTGTGTTATTTGGGGGGAAGTCTGTACAACATCTTGA 782
 Db 539 sLeuAlaPheSerTyrMetAsePaspValValLeuGlyAlaLysSerValGlnHisLeuGl 559
 QY 783 GTCCCTTTTACTCTATTACCAATTTCTTTTGTCTTTTGGGTATATCATTTAAACCTTAA 842
 Db 559 uSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAenProAs 579
 QY 843 TAAACCAAAACGTTGGGGCTACTCCCTTAACTTCATGGGATATGTAAATTTGGAAGTTGGGG 902
 Db 579 nLysThrLysArgTrpGlyTyrSerLeuAenPheMetGlyTyrValIleGlySerTrpGl 599
 QY 903 TACTTTACCGCAGGACATATTGTACTAAACCTCAAGCAATGTTTTTCGAAAACTGCTGT 962
 Db 599 yThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuProVa 619
 QY 963 AAATAGACCTATTGATTGGAAAGTATGTCAAAGAATTTGGGGTCTTTTGGGCTTTTCTGTC 1022
 Db 619 lAsnArgProIleAseTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAlaAl 639
 QY 1023 CCCTTTTACACAAATGGGCTATCCGCTTGTATGCTTATATGATGATATACAACTAA 1082

Db 639 aProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSerIy 659
Qy 1083 GCAGGCTTTCACTTTCTCGCAACTTACAAAGGCTTTCTGTGTAACAATATCTGAACCT 1142
Db 659 sGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysGlnGlnTyrLeuHisLe 679
Qy 1143 TTAACCGCTTCCCGCGGCAAGCTCTCGCTCTCCGAAGTGTGTGCTGACGACACCCCGAC 1202
Db 679 uTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrProth 699
Qy 1203 TGGATGGGCTTGGCCATAGCCATAGCGATGCGCTGCGAAGCTTTCTGCTCTCTGCTGCC 1262
Db 699 xGlyTyrGlyLeuAlaIleGlyHisArgMetArgGlyThrPheValAlaProLeuPr 719
Qy 1263 GATCCATATCGCGAACTCTAGCAGCTTGTGTGCTCGAGCGGTCTGGAGCAAACT 1322
Db 719 oIleHisThrAlaGlnLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaIysLe 739
Qy 1323 TATCGGAACGACAACTCTGCTCTCTCTCTCGGAATACACCTCTCTTCCATGGCTGCT 1382
Db 739 uIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProThrLeuLe 759
Qy 1383 AGGTGTGCTGCCAACTGAGTCTCTGCGCGGAGCTCTTGTCTTACGTCCTCGCTCGCGCT 1442
Db 759 uGlyCysAlaAlaAsnTrpIleuArgGlyThrSerPheValTyrValProSerAlaLe 779
Qy 1443 GAATCCCGCGGACGACCGCTCTCGGCGCGTGTGGGCTCTTACCGTCCCTCTTCTCATCT 1502
Db 779 uAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHisLe 799
Qy 1503 GCGGTCGCGCGGACGCGGCGGCGACCTCTCTTACGCGGTCTCCCGGTATGCGCTTC 1562
Db 799 uProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValProSe 819
Qy 1563 TCATCTCGCGGACGCTGCACTTCTGCTTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
Db 819 rHisLeuProSerArgValHisPheAlaSerProLeuHisValAlaTrpArgProPro 838

RESULT 15
O42041_HPBVO PRELIMINARY; PRT; 843 AA.
AC O42041_1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE .DNA polymerase.
OS Hepatitis B virus.
OC Viruses; Retroviral viruses; Hepadnaviridae; Orthohepadnavirus.
ON NCBI_TaxID=10407;
RX NUCLEOTIDE SEQUENCE.
RL MEDLINE=97160188; PubMed=9007704; DOI=10.1016/S0168-8278(96)80280-7;
RA Aashina Y., Enomoto N., Kurotsaki M., Sakuma I., Izumi N., Marumo F.,
Sato C.;
RT "Sequential changes in full-length genomes of hepatitis B virus
accompanying acute exacerbation of chronic hepatitis B";
J. Hepatol. 25:787-794(1996).
DR EMBL; D50518; BAA23442.1; -; Genomic DNA.
DR EMBL; D50517; BAA23435.1; -; Genomic DNA.
DR GO; GO:0003677; P:DNA binding; IEA.
DR GO; GO:0003887; P:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0004519; P:endonuclease activity; IEA.
DR GO; GO:0016787; P:hydrolyase activity; IEA.
DR GO; GO:0004523; P:ribonuclease H activity; IEA.
DR GO; GO:0003723; P:RNA binding; IEA.
DR GO; GO:0003964; P:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001462; DNaPol_viral_C.
DR InterPro; IPR00201; DNaPol_viral_N.
DR InterPro; IPR000477; RVase.
DR Pfam; PF00336; DNaPol_viral_C; 1.
DR Pfam; PF00242; DNaPol_viral_N; 1.

DR Pfam; PF00078; RVT_1; 2.
DR ProDom; PD000814; DNaPol_viral_C; 1.
SQ SEQUENCE 843 AA; 94462 MW; 1BCCF79639BB140D CRC64;
Alignment Scores: 2.79e-200 Length: 843
Pred. No.: 2802.00 Matches: 525
Score: 98.52% Conservative: 8
Percent Similarity: 97.04% Mismatches: 7
Best Local Similarity: 47.44% Indels: 2
Query Match: 2 Gaps: 0
DB: US-10-761-006a-1 (1-3215) x O42041_HPBVO (1-843)

Qy 1 CTCCACAACATTCCACCAGCTCTGCTAGATCCAGGCTGAGGGCTATATTTTCTCTGC 60
Db 304 LeuHisAsnIleProProSerAlaArgSerGlnSerGlnGlyProIlePheSerCys 323
Qy 61 TGTGTGCTCCAGTTCCGGAACTAAACCTGTTCCGACTTACTGCTCTCCCATATCGTC 120
Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 343
Qy 121 AATCTTCTCGAGACTTGGGGACCTCGACCGAACATGAGAACACCAATCAGGATTCCT 180
Db 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
Qy 181 AGACCCCTGCTGCTGTTTACAGCGGGGTTTTCTGTTGACAAAGAACTCTCAATACC 240
Db 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
Qy 241 GCAGACTCTAGACTCTG-GTGGACTTCTCTCAATTTCTAGGGGGAGCACCCAGCTGTT 299
Db 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisIleVal -Se 403
Qy 300 CTGCGCAAAATTCGAGTCTCCCAACTCCAACTCACTCACTCACTCACTCTGTTCTCCAAATTT 359
Db 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
Qy 360 GTCTGCTGCTATCGCTGATGTGCTGCGGGTTTTATCATATTTCTTCTTCTTCTGCTGC 419
Db 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
Qy 420 TATGCTCTCATCTTCTTGTGTTCTCTGACTTACCAAGGTATGTTGCCGTTGTTCTCCT 479
Db 443 aMetProHisLeuLeuValGlySerGlyLeuProArgTyrValAlaArgLeuSerSe 463
Qy 480 TACTTCCAGGAACATCAACCAACGAGCAGCGGGGCATGCAAGACTGCACTGCTCTCTCCTC 539
Db 463 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisGluSerCysSe 483
Qy 540 AAGGAAACTTACGTTTCTCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 599
Db 483 xArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 503
Qy 600 GTATTCCATCCCATCATCTGCGGCTTCCAGAGATTCCTATGGAGTGGGCTTCTCCTCC 659
Db 503 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
Qy 660 GTTCTCTGCTGCTCAGTTTACTAGTGCATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 719
Db 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi 543
Qy 720 CTGTTTGGCTTTTCACTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 779
Db 543 sCysLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLe 563
Qy 780 TGAGTCCCTTTTACCTCTTATACCAATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 839
Db 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
Qy 840 TAATAAACAACAGTTGGGGCTACTCCCTTAACCTTCACTGATATGTAATGGAAGTTG 899
Db 583 oAsnLysThrLysArgTrpGlyThrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603

QY 900 GGGTACTTTTACCGCAGGAACATATGTACTATAAACTCAAGCAATGTTTTCGAAAACTGCC 959
Db |||||||
QY 603 pGlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPr 623
Db |||||||
QY 960 TCTAATAGACCTATTGATGTGAAAGTATGTCAAAGAATTGTGGGTCTTTTGGGCTTTGC 1019
Db |||||||
QY 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
Db |||||||
QY 1020 TCGCCCTTTTACACATGTGGCTATCTCCCTGATGCTTTATATGATGATGATATCAATC 1079
Db |||||||
QY 643 aAlaProPheThrGlnCysGlyTrpProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
Db |||||||
QY 1080 TAAGCAGGGCTTTCACCTTTCTCCCAACTTACAAAGCGCTTCTGTGTAAACAATATCTGAA 1139
Db |||||||
QY 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysGlnGlnTyrLeuHi 683
Db |||||||
QY 1140 CCTTTACCCCGTTGCCGCAACGGTCCGCTCTCTGCCAAGTGTGTGTGACGCAACCCC 1199
Db |||||||
QY 683 sLeuTyrProValAlaArgGlnArgSerGlyValCysGlnValPheAlaAspAlaThrPr 703
Db |||||||
QY 1200 CACTGGATGGCGCTGGCCATAGGCCATCAGCGCATGGCTGGAACCTTTCTGGCTCCCT 1259
Db |||||||
QY 703 oThrGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProle 723
Db |||||||
QY 1260 GCGGATCCCATCTGCGGAACCTCTAGCAGCTGTGTTTGTCTCGCAGCGGTCTGGAGCAA 1319
Db |||||||
QY 723 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLy 743
Db |||||||
QY 1320 ACTTATCGGAACCGCAACTCTGTGTCTCTCTCGGAAATACACCTCTTTCCATGGCT 1379
Db |||||||
QY 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProtriple 763
Db |||||||
QY 1380 GCTAGGGTGTGTCGCAACTGGATCCTGCGCGGAGCTCTTTGCTAGCTCCGCTCGGC 1439
Db |||||||
QY 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
Db |||||||
QY 1440 GCTGAATCCCGGAGCGACCGCTCTCGGGCGGTGTTGGGGCTCTACCGTCCCTCTTCA 1499
Db |||||||
QY 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 803
Db |||||||
QY 1500 TCTGCGGTTCCGCGCACCAACCGGCGCACCTCTCTTTACGCGGTCTCCCGTATGTGCC 1559
Db |||||||
QY 803 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
Db |||||||
QY 1560 TTCTCATCTGCGGACCGGTGTCACCTTCGCTTCACCTTCGACGTCGATGGAGACCACC 1619
Db |||||||
QY 823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 843
Db |||||||
QY 1620 G 1620
Db ||
QY 843 O 843

Search completed: December 1, 2005, 23:59:04
Job time : 772.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 1, 2005, 23:15:45 ; Search time 106 Seconds

(without alignments)
5836.545 Million cell updates/sec

Title: US-10-761-006A-1

Perfect score: 5907

Sequence: 1 CTCACACATTCACACAG.....CCTCAGGCCACGAGTGGAA 3215

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delxet 7.0

Searched: 283416 segs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/csm2_1/USPTO.spool_p/US10761006/runat_01122005_113953_26882/app.query.fasta_1.3399
-DB=PIR -OPT=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODES=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10761006 @CGN 1 1 163 @runat_01122005_113953_26882 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR 80.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2807	47.5	843	1 JDVLVR	DNA-directed DNA p
2	2789	47.2	843	2 S35527	DNA-directed DNA p
3	2750	46.6	843	1 S43491	DNA-directed DNA p
4	2734.5	46.3	842	1 JDVLVS	DNA-directed DNA p
5	2688	45.5	825	2 T13473	DNA-directed DNA p
6	2678.5	45.3	827	2 T13468	DNA-directed DNA p
7	2654	44.9	832	2 S71785	DNA-directed DNA p
8	2652	44.9	832	1 JDVLVB	DNA-directed DNA p
9	2651	44.9	832	1 S47406	DNA-directed DNA p
10	2643	44.7	832	1 JDVLVA	DNA-directed DNA p
11	2643	44.7	843	1 JDVLJ3	DNA-directed DNA p
12	2642	44.7	832	1 JDVLAI	DNA-directed DNA p
13	2637	44.6	832	1 S20757	DNA-directed DNA p
14	2628	44.5	832	2 S67505	DNA-directed DNA p

15	2628	44.5	843	1 JDVLJ1	DNA-directed DNA p
16	2623	44.4	832	1 JDVLCP	DNA-directed DNA p
17	2610	44.2	843	1 JDVLJ2	DNA-directed DNA p
18	2608	44.2	845	1 JDVLKS	DNA-directed DNA p
19	2607	44.1	832	1 S20752	DNA-directed DNA p
20	2607	44.1	845	1 JDVLVD	DNA-directed DNA p
21	2582	43.7	843	1 JQ2229	DNA-directed DNA p
22	2230	37.8	750	1 JDVLVH	gene X/C fusion pr
23	1792	30.3	359	2 S04570	DNA-directed DNA p
24	1719	29.1	884	1 JDVL7	DNA-directed DNA p
25	1716	29.1	884	1 JDVL59	DNA-directed DNA p
26	1714	29.0	556	1 JDVL64	DNA-directed DNA p
27	1708	28.9	884	1 JDVLW8	DNA-directed DNA p
28	1698	28.7	879	1 JDVLC	DNA-directed DNA p
29	1668	28.2	881	1 JDVLS	DNA-directed DNA p
30	1649	27.9	883	1 JDVLC2	large surface anti
31	1472	24.9	400	1 SAVLA	surface antigen -
32	1452	24.6	400	2 S35528	surface antigen -
33	1391.5	23.6	445	2 S43492	large surface anti
34	1380	23.4	389	1 SAVLCP	large surface anti
35	1376	23.3	389	1 SAVLBH	large surface anti
36	1372	23.2	389	2 S67506	surface antigen -
37	1369	23.2	389	2 S20745	major surface anti
38	1368	23.2	400	1 JQ1575	large surface anti
39	1367	23.1	389	2 S32202	surface antigen -
40	1363	23.1	389	2 S47407	surface antigen -
41	1363	23.1	389	2 S20753	surface antigen -
42	1361	23.0	400	1 SAVLVD	surface antigen -
43	1356	23.0	389	2 S20749	large surface anti
44	1355	22.9	384	2 T13469	large surface anti
45	1354	22.9	389	1 SAVLAH	large surface anti

ALIGNMENTS

RESULT 1

JDVLVR

DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype adr)

C/Species: hepatitis B virus, HBV

C/Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 09-Jul-2004

C/Accession: A00704

R/Ono, Y.; Onda, H.; Sasada, R.; Igarashi, K.; Sugino, Y.; Nishioka, K.

Nucleic Acids Res. 11, 1747-1757, 1983

A/Title: The complete nucleotide sequences of the cloned hepatitis B virus DNA; subtype

A/Reference number: A93460; MUID:83168919; PMID:6300776

A/Accession: A00704

A/Molecule type: DNA

A/Residues: 1-843 <ONO>

A/Cross-references: UNIPROT:P03157; UNIPARC:UPI00001297FE

C/Superfamily: hepatitis virus DNA-directed DNA polymerase

C/Keywords: DNA biosynthesis; nucleotidyltransferase

Alignment Scores:

Pred. No.:	1.13e-199	Length:	843
Score:	2807.00	Matches:	528
Percent Similarity:	98.33%	Conservative:	3
Best Local Similarity:	97.78%	Mismatches:	8
Query Match:	47.52%	Indels:	2
DB:	1	Gaps:	0

US-10-761-006A-1 (1-3215) x JDVLVR (1-843)

QY	4	CACAACTTCCACCAAGCTCTGTAGATCCAGGCTGAGGGCCCTATATTTCTCTGG	63
Db	305	HisAenilleProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCysTrp	324
QY	64	TGGCTCCAGTTCGGACAGTAACCCCTGTCGACTACTCCCTCTCCATATCTCAAT	123
Db	325	TrpLeuGlnPheArgAsnSerlyseProCysSerAspTyrCysLeuThrHisIleValAen	344
QY	124	CTTCTCGAGGACTGGGGACCCCTGACACCAATGAGAACACAACTCAGGATTCCTAGG	183
Db	345	LeuLeuGluAspTrpGlyProCysThrGluHisGlyHisAsnIleArgIleProArg	364

QY 241 GCAGAGTCTAGACTCTG-CTGGACTTCTCAATTTCTAGGGGAGCACCCAGTGTTC 299
DB 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
QY 300 CTGGGCAAAATTCGCGAGTCCCAACCTCCAACTCACTCAACCAACCTTCTGTCTCCAAATTT 359
DB 403 rTrpProlysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
QY 360 CTCTGGGTATCGCTGGAGTGTCTGCGCGGTTTATCATATTCCTTCACTTCCTGCTGC 419
DB 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyHisIleProLeuHisProAlaAl 443
QY 420 TATGCTCATCTTCTGTGTGGTCTCTCGGACTACCAAGTATGTGGCCGTTCTGCTCCTC 479
DB 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyValAlaAlaArgLeuSerSe 463
QY 480 TACTTCCAGGAACATCAACACACGAGCGGCCCATGCAAGACCTTCGACGACTCCTGCTC 539
DB 463 rThrSerArgAsnIleAsnTyHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
QY 540 AAGGAACTCTACGTTTCCCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 599
DB 483 rArgAsnLeuTyValSerLeuLeuValTyHisThrPheGlyArgLeuHisLe 503
QY 600 GTATTCCCATCCATCATCTCTGGCTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCC 659
DB 503 uTySerHisProIleLeuGlyPheArgTyHisLeProMetGlyValGlyLeuSerPr 523
QY 660 GTTCTCTCGGCTCAGTTTACTAGTGCATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 719
DB 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHis 543
QY 720 CTGTTTGGCTTTCAGTTATATGATGATGTGTATGTGGGGCGAAGTCTGTCAACATCT 779
DB 543 sCysLeuAlaPheSerTyMetAspValValLeuGlyAlaLysSerValGlnHisLe 563
QY 780 TGAGTCCCTTTTACCTTATTACCAATTTCTTTTGTCTTTTGTCTTTTGTCTTTTGTCTTT 839
DB 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
QY 840 TAATAAAACCAACGTTGGGCTACTCTTAACTTCATGGATGATGTAATTTGGAAGTTG 899
DB 583 oAsnLysThrLysArgTrpGlyTySerLeuAsnPheMetGlyTyValIleGlySerTr 603
QY 900 GGGTACTTTACGCGAGGAACATATTGTACTAAACTCAAGCAATGTTTTCGAAAACCTGCC 959
DB 603 pGlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPr 623
QY 960 TGTAAATAGACCTATTGATGGAAAGTATGTCAAGAATTGGGCTTTTTCGGCTTTGC 1019
DB 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
QY 1020 TSCCCCTTTTACCAATGTGGCTATCTGCTGTATGCTTATATGATGATGATATCAATC 1079
DB 643 aAlaProPheThrGlnCysGlyTyProAlaLeuMetProLeuTyAlaCysIleGlnSe 663
QY 1080 TAAGCAGGCTTTCACTTTCTCGCAACTTACAGGCTTTCTGTGTAAACAATATCTGAA 1139
DB 663 rLysGlnAlaPheThrPheSerProThrTyHisAlaPheLeuCysGlnGlnTyHisLeuHi 683
QY 1140 CTTTACCCGTTGCCCGCAACGTTCCGTTCTCTCCCAAGTGTGTGCTGCAACACCCC 1199
DB 683 sLeuTyProValAlaArgGlnArgSerGlyLeuCysGlnValPheGlyAspAlaThrPr 703
QY 1200 CACTGTAGTGGGCTTCGCCCATAGCCATGCGCTGCTGGAACCTTTCTGGCTCCTCT 1259
DB 703 oThrGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLe 723
QY 1260 GCCGATCCATPACTCGCGAACTCTAGCAGCTTCTTTTGTCTCGCAGCGGCTCTCGAGCAA 1319
DB 723 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaTy 743
QY 1320 ACTTATCGGAACCGCAACTCTGTTGTCTCTCTCGGAATAATACCTCTCTTCCATGGCT 1379

DB 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyThrSerPheProTrpLe 763
QY 1380 GCTAGGTGTGCTGCCAACATGATCTCTGCGGGGAGACGTCTCTTGTCTACGTCCCGTCGGC 1439
DB 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyValProSerAl 783
QY 1440 GCTCAATCCCGCGAGGACCGCTCTCGGGGCGGCTTGGGGCTCTACCGTCCCTCTCTTCA 1499
DB 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyArgProLeuLeuHi 803
QY 1500 TCTCGCTTCGCGCGCGACACCGCGCGCACCTCTCTTTACCGGCTCTCCCGTATGTGCC 1559
DB 803 sLeuProPheGlnProThrThrGlyArgThrSerLeuTyAlaValSerProSerValPr 823
QY 1560 TTCTCATCTCGCGAGCGGTGTGCATCTTCGCTTACCTCTCGCAGCTCGCATGAGACCACC 1619
DB 823 oSerHisLeuProValArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 843
QY 1620 G 1620
DB 843 o 843
RESULT 3
S43491
DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype adr)
C:Species: hepatitis B virus, HBV
A:Variety: subtype adr
C:Date: 07-Sep-1994 #sequence_revision 22-Nov-1996 #text_change 09-Jul-2004
C:Accession: S43491
R:Loncarevic, I.F.; Zentgraf, H.; Schroeder, C.H.
Nucleic Acids Res. 18, 4940, 1990
A:Title: Sequence of a replication competent hepatitis B virus genome with a preX open
A:Reference number: S12598; MUID:90370503; PMID:2395664
A:Accession: S43491
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-843 <LON>
A:Cross-references: UNIPROT:Q67952; UNIPARC:UPI00000EE7B3; EMBL:X52939; NID:9457780; PFI
A:Experimental source: subtype adr
C:Genetics:
A:Gene: P; pol
C:Superfamily: hepatitis virus DNA-directed DNA polymerase
C:Keywords: DNA biosynthesis; nucleotidyltransferase
Alignment Scores:
Pred. No.: 1.89e-195 Length: 843
Score: 2750.00 Matches: 517
Percent Similarity: 97.04% Conservative: 8
Best Local Similarity: 95.56% Mismatches: 15
Query Match: 46.55% Indels: 2
DB: 1 Gaps: 0
US-10-761-006A-1 (1-3215) x S43491 (1-843)
QY 1 CTCGACACATTCACCAAGCTCTCTAGATCCAGGTCGAGGGGCTATATTTCTCTC 60
DB 304 LeuHisAsnIleProProSerSerAlaArgSerLysSerGlyProLeuPheProCys 323
QY 61 TGGTGGCTCCAGTTCGGAAACAGTAAACCTGTCCTCGACTACTCTCCCATATATCGTC 120
DB 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyCysLeuThrHisIleVal 343
QY 121 AATCTTCTCAGGACTGGGGACCTTCGACCGAACATGGAGAACACAAATCAGGATTCCT 180
DB 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
QY 181 AGACCCCTGCTCGTGTATACAGCGGGGTTTTCTCGTTGACAGAAGATCCTCAATACC 240
DB 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
QY 241 GCAGAGTCTAGACTCTG-CTGGACTTCTCTCAATTTCTAGGGGAGCACCCAGTGTTC 299

```
Db      384  ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnThrGlnVal-Se 403
Qy      300  CTGGCCAAAATTGCGAGTCCCAACCTCCAATCACTCAACCACTCTTGCTCCTCCAAATTT 359
Db      403  rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
Qy      360  GTCTGGCTATCGCTGGATGTCTGCGGCGTTTATCATATTCCTCTTCATCTGCTGTC 419
Db      423  uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
Qy      420  TATGCTCATCTCTTGTGTCTCTCGGACTACCAAGGTATGTGCGCGTGTGTCCTC 479
Db      443  aMetProHisLeuLeuValGlySerSerGlyLeuGlnArgTyrValAlaAlaArgLeuSerSe 463
Qy      480  TACTTCCAGGAACATCAACACCAACGCGGCGCCATGCAAGACCTTCGACGAACTGCACTT 539
Db      463  rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
Qy      540  AAGGAAACTCTACGTTCCCTCTGTTGCTGTACAAAACCTTCGACGAACTGCACTT 599
Db      483  rLysHisLeuTyrValSerLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 503
Qy      600  GTATTCCCATCCCATCCCTGGGCTTTCGCAAGATTCCTATGGAGTGGGCGCTCAGTCC 659
Db      503  uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
Qy      660  GTTTCCTCTGCTCAGTTTACTAGTGCATTTGTTCAGTGGTTCGTAGGGCTTTCCCCCA 719
Db      523  oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValCysArgAlaPheProHi 543
Qy      720  CTGTTGGCTTTCAGTTATATGGATGATGTGGTATTTGGGGCGAAGCTGTACAAATCT 779
Db      543  sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
Qy      780  TGAGTCCCTTTTACCTCTATTACCAATTTCTTTGTCTTTGGTATACATTTAAACCC 839
Db      563  uGluSerLeuPheThrSerIleThrAsnPheMetCysLeuSerLeuGlyIleHisLeuAsnPr 583
Qy      840  TAATAAAACCAACCGTTGGGGCTACTCCCTTAACCTTCATGGGATATGTAATTGGAAGTTG 899
Db      583  oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
Qy      900  GGGTACTTTACCGCAGGACATATGTACTAAAACCTCAAGCAATGTTTTCGAAAACCTGCC 959
Db      603  pGlyThrLeuProGlnGluHisIleValGlnLysLeuLysGlnCysPheArgLysLeuPr 623
Qy      960  TGTAAATAGACTATTGATTGGAAGTATGTCAAAGAATTGTGGGTCTTTTGGGCTTGC 1019
Db      623  oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
Qy      1020  TGCCCTTTTACAAATGTGGCTATTCCTGCTTGTATGCTCTTTATATGATGATATACAAATC 1079
Db      643  aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
Qy      1080  TAACGAGCTTTCACTTCTCGCAACTTACAGCGCTTCTGTGTAAACAATATCTGAA 1139
Db      663  rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuLeuCysLysGlnTyrLeuAs 683
Qy      1140  CTTTACCCCTGTCGCGGCAACGCTCGGTCTCTGCCAAGTGTTCGTGACGCAACCCC 1199
Db      683  nLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
Qy      1200  CACTGGATGGGGCTTGGCCATAGCCCATCGCGCATGGCTGGAACTTTCTGGGCTCCTCT 1259
Db      703  oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValAlaProLe 723
Qy      1260  GCCCATCCACTCTGGGAACCTCTAGCAGCTTGTGTTGCTCGCAGCGCGTCTGAGCAAA 1319
Db      723  uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaTh 743
Qy      1320  ACTTATCGGAACCGACAACTCTGTTGTCTCTCTCGGAATAATACACCTCTCTTCCATGCT 1379
Db      743  rLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 763
```

```
Qy      1380  GCTAGGTGTCTCTCCCAACTGGATCTTCGCGGGACGTCCTTTGTCTACGTCCCGTCGGC 1439
Db      763  uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
Qy      1440  GCTGAATCCCGCGGACGACCGCTCTCGGGGCGGTTTGGGGCTCTACCGTCCCTCTTCA 1499
Db      783  aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuAr 803
Qy      1500  TCTGCGGTTCCGGCGCAGCACCGGCGCACCTCTCTTTACGCGGTCTCCCGTATGTGCC 1559
Db      803  gLeuSerPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
Qy      1560  TTCTCATCTCGCGGACCGTGTGCATCTTCGCTTCACCTCTGCACCTCGCATGGAGCACACC 1619
Db      823  oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaIleTrpArgProPr 843
Qy      1620  G 1620
Db      843  o 843

RESULT 4
JDVLVS
DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype adr, mutant)
C:Species: hepatitis B virus, HBV
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C:Accession: S04568
R:Rho, H.M.; Kim, K.; Hyun, S.W.; Kim, Y.S.
Nucleic Acids Res. 17, 2124, 1989
A:Title: The nucleotide sequence and reading frames of a mutant hepatitis B virus subtyp
A:Reference number: S04568; MUID:89183619; PMID:2928116
A:Accession: S04568
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-842 <RHO>
A:Cross-references: UNIPROT:P31870; UNIPARC:UPI00001297FB; EMBL:X14193
C:Superfamily: hepatitis virus DNA-directed DNA polymerase
C:Keywords: DNA biosynthesis; nucleotidyltransferase

Alignment Scores:
Pred. No.:      2,65e-194      Length:      842
Score:          2734.50      Matches:    517
Percent Similarity: 96.86%      Conservative: 7
Best Local Similarity: 95.56%      Mismatches: 15
Query Match:     46.29%      Indels:     3
DB:              1              Gaps:       1

US-10-761-006A-1 (1-3215) x JDVLVS (1-842)
Qy      1  CTCACAACATTCACCAAGCTCTGTAGATCCAGGGTGAGGGGCTATATTTTCCTGC 60
Db      304  LeuHisIleSerProSerProAlaArgSerGlnSerGluGlyProIlePheSerSer 323
Qy      61  TGTGGCTCAGTTCGGGAACAGTAAACCCCTGTTCCGACTACTGCTCTCCCATATCGTC 120
Db      324  TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 343
Qy      121  AATCTCTTCGAGGACTGGGACCCCTGCACCGACATCGAGAACACACATCAGGATTCCT 180
Db      344  AsnLeuLeuGlnAspTrpGlyProCysThrGluHisGlyLeuHisAsnIleArgIlePro 363
Qy      181  AGGACCCCTCTCTGTTTACAGCGGGGTTTTCTCGTTGACAAGAAATCCTCACAATACC 240
Db      364  ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
Qy      241  GCAGAGTCTAGACTCTG- GTGGACTTCTCTCAATTTTCTAGGGGGAGACACCCAGTGTTC 299
Db      384  ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
Qy      300  CTGGCCAAAATTGCGAGTCCCAACCTCCAATCACTCAACCACTCTTGCTCCTCCAAATTT 359
Db      403  rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
```


360 GTCTGGCTACGCTGGATGTCTGGCGGTTTATCATATTCTCTTCATCTCTGCTGC 419
423 userftrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
420 TATGCTCTCATCTCTTGTGTGTTCTTCTGGACATCAAGGTATGTTCCCGTTTGTCTTC 479
443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaargLeuSerSe 463
480 TACTTCAGGAACTACACCAACAGCAGCGGGCCATGCAAGACTGACAGCTCTCTCTC 539
463 rThrSerArgAsnIleAsnHisGlnHisGlyThrMetGlnAspLeuHisaspSerCysSe 483
540 AAGGAAACTCTACCTTCCCTCTGTTGCTGTACAAACCTTCGACGGAACCTGCACCTT 599
483 rArgAsnLeuTyrValSerLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 503
600 GTATTCCCATCCCATCTCTGGGCTTTCGCAAGATTCTTATGGAGTGGGCTTCAGTCC 659
503 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyGlyGlyLeuSerPr 523
660 GTTTCCTCTGCTCAGTTTACTAGTCCATTGTTTCAGTGTGTTCTGAGGCTTTCCTCCCA 719
523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 543
720 CTGTTTGGCTTTCAGTTTATATGATGATGTTGTTGGGCGGAGTCTGTACACATCT 779
543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
780 TGAGTCCCTTTTACCTATTACCAATTTCTTCTTGTGTTGGTATACATTTAAACCC 839
563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
840 TAATAAAACCAAGCTTGGGCTACTCCCTTAACTTCATGGATATGTAATTTGGAAGTTG 899
583 oAsnLysThrLysArgTyrGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
900 GGGTACTTTACCGAGGACATATGTACTAAACTCAAGCAATGTTTTCGAAACCTGCC 959
603 pGlyThrLeuProGlnGluHisIleValLeuLysIleLysGlnCysPheArgLysLeuPr 623
960 TGTAATAAGACTATTGATTCGAAAGATATGTCAAGAAATCTGGTGTCTTTGGCTTTC 1019
623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
1020 TGCCCTTTTACAAATGTGCTACTCTGCTCTGATGCTTATATGATGATATACATC 1079
643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
1080 TAAGAGGCTTTCTCTTCGCAACTTACAGGCTTTCTGTGTAAACAAATATCTGAA 1139
663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuLeuCysLysGlnTyrLeuHi 683
1140 CTTTACCCCTTCCCGGCAACCGTTCGGTCTCTGCAAGTGTTCGTGACGCAACCCC 1199
683 sLeuTyrProValAlaArg---ArgThrAlaLeuCysGlnValPheAlaAspAlaThrPr 702
1200 CACTGTGATGGGCTTGGCCATAGCCATCAGCGATGCTGGAACCTTTCGGTCTCTCT 1259
702 oThrGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLe 722
1260 GCGGATCCATCTCGGAACTCTTAGCAGCTTGTGTTGCTCGCAGCCGGTCTGGAGCAAA 1319
722 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys 742
1320 ACTTATCGGAACCGCAACTCTGTGCTCTCTCGGAAATACACCTCTTTCATGGCT 1379
742 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 762
1380 GCTAGGGTGTCTGCCAACTGATCTCGCGGGAGCTCTTGTGTACGTCCTCGGCTGGC 1439
762 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrTyrPheValTyrValProSerAl 782
1440 GCTGAATCCCGGACGACCGGCTCTCGGGGGCGGTTTGGGGCTCTACCGTCCCTCTCTCA 1499

782 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuIleargProLeuLeuHi 802
1500 TCTGCGCTTCCGCGCCACACGCGGCGCACCTCTCTTTAGCGGTCTCCCGGTATGTGCC 1559
802 sLeuArgPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 822
1560 TTTCTCATCTGCGGACCGGTGTCACCTTCGCTTTCACCTCTGACGTGCGATGGAGACCC 1619
822 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 842
1620 G 1620
842 o 842
RESULT 5
T13473
DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (isolate 29Y11HCC)
C:Species: hepatitis B virus, HBV
A:Variety: isolate 29Y11HCC
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C:Accession: T13473
R:Takahashi, K.; Akahane, Y.; Hino, K.; Ohta, Y.; Mishiro, S.
Arch. Virol. 143, 2313-2326, 1998
A:Title: Hepatitis B virus genomic sequence in the circulation of hepatocellular carcin
A:Reference number: Z17684; MUID:99129050; PMID:9930189
A:Accession: T13473
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-825 <TAK>
A:Cross-references: UNIPROT:Q9Y2S3; UNIPARC:UPI00000F4BF5; EMBL:AB014388; NID:G3582381;
A:Experimental source: Japanese patient with hepatocellular carcinoma isolate 29Y11HCC
C:Genetics:
A:Gene: P
A:Introns: 303/3
C:Superfamily: hepatitis virus DNA-directed DNA polymerase
C:Keywords: DNA biosynthesis; nucleotidyltransferase
Alignment Scores:
Pred. No.: 7.33e-191 Length: 825
Score: 2688.00 Matches: 504
Percent Similarity: 95.92% Conservative: 13
Best Local Similarity: 93.51% Mismatches: 21
Query Match: 45.51% Indels: 2
DB: Gaps: 2
US-10-761-006A-1 (1-3215) x T13473 (1-825)
QY 7 AACATTCACCAAGCTCTGCTAGATCCAGGGTGAGGGGCTATATTTCTCTGCTGTGG 66
288 HisLeuSerThrThrLysArgGlnSerSerSerGlyHisAlaValGluThrCysTrpTrp 307
67 CTCAGTTCGGACAGTAACCTGTTCCGACTACTCTCTCCATATCGTCAATCTT 126
308 LeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleValAsnLeu 327
127 CTCGAGACTGGGACCTCCACCGAATATGGAGAACCAACATCAGATTCCTAGGACC 186
328 LeuGluAspTrpGlyProCysThrGluHisGlyGluHisIleAargIleProArgThr 347
187 CTTGCTCGTGTATAGCGGGGTTTCTCGTTGACAAAGATCTCTCAATACGACGAG 246
348 ProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThrThrGlu 367
247 TCTAGACTCTG-GTGGACTTCTCTCAATTTCTAGGGGGAGCACCCAGCTTCTCTGCC 305
368 Ser***LeuValValAspPheSerGlnPheSerArgGlySerThrGlnVal-SerTrpPr 387
306 AAAATTCCGAGTCCCAACCTCCAACTCACCTCACCAAGCTCTTGTCTCCCAATTTGCTCG 365
387 oLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeuSerTr 407
366 GCTATCTGCGATGTGTCTGCGCGGTTTTATCATATTTCTCTTCTCATCTCTGCTATGCC 425

```

Db      407 pLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAlaMetPr 427
      |||
QY      426 TCATCTCTTGTGTTCTTCTCGACTACCAAGGTATGTGGCGTGTCTCTACTTTC 485
      |||
Db      427 oHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSerThrSe 447
      |||
QY      486 CAGGAACATCAACACACGAGCGGGCCATGCAAGACCTGCAAGCTCTCTGCTCAAGGAA 545
      |||
Db      447 rArgAenIleAen**GinHisGlyThrMetGlnAspLeuHisAspSerCysSerLysHi 467
      |||
QY      546 ACTCTACGTTTCCCTCTGTTCTCTGACACAAACCTTCGAGCGGAACCTGCTATTC 605
      |||
Db      467 sLeuTyrValSerLeuLeuLeuTyrIlePheGlyArgLysLeuHisLeuTyrSe 487
      |||
QY      606 CCATCCCATCATCTCTGGGCTTTCGAAAGTCTCTATGGGAGTGGGCTCAGTCCGTTCT 665
      |||
Db      487 rHisProIle**LeuGlyPheArgLysIleProMetGlyValGlyLeuSerProPheLe 507
      |||
QY      666 CTGCTCTCAGTTTACTAGTGCATTTGTTTCACTGTTCTGAGGGCTTTTCCCTACTGTTT 725
      |||
Db      507 uLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHisCysLe 527
      |||
QY      726 GCGTTTCACTTATATGATGATGTGGTATTTGGGGGGAAGTCTGTACAACTTTGAGTC 785
      |||
Db      527 uAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLeuGluAl 547
      |||
QY      786 CTTTCTTACCTCTATTACCAATTTCTTTGCTTTGCTTGGGTATACATTTAAACCTTAATAA 845
      |||
Db      547 aLeuTyrThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnProAsnL 567
      |||
QY      846 AACCAAACTTGGGGCTACTCCCTTAACTTCACTGGGATATGTAATTTGGAAGTTGGGGTAC 905
      |||
Db      567 sThrLysArgGTPGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTrpGlyTh 587
      |||
QY      906 TTTACGGCAGGAACATATTGTTACTAAACTCAAGCAATGTTTTCGAAACCTCGCTGTAAA 965
      |||
Db      587 rLeuProGlnAspHisIleValGlnLysLeuLysGlnCysPheArgLysLeuProValAs 607
      |||
QY      966 TAGACCTATTGATTGGAAGTATGCAAGAATTGTTGGGTCTTTTGGGCTTTGCTGCCCC 1025
      |||
Db      607 nArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAlaAlaPr 627
      |||
QY      1026 TTTTACCAATGTGGCTACTCTGCTTGTATGCTTTTATATGCAATGATATCAATCTAAGCA 1085
      |||
Db      627 oPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSerLysGl 647
      |||
QY      1086 GGTCTTCACTTCTCGCCAACTTACAAAGGCTTCTGTGTAAACATATCTGAACCTTTA 1145
      |||
Db      647 nAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAsnLeuTy 667
      |||
QY      1146 CCGGTTTCCCGCAACGGTCCGGTCTCTGCCAAGTGTTCGTGACGCAACCCCACTG 1205
      |||
Db      667 rProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrProThrGl 687
      |||
QY      1206 ATGGGGCTTGGCCATAGGCGCATCAGCGCATGGCTGGAACTTTCTGGCTCTCTGCGCGAT 1265
      |||
Db      687 yTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLeuProIl 707
      |||
QY      1266 CCATATCGGGAACCTCTCTAGCAGCTTGTGTTTGTCTCGACCGGCTGTGGAGCAAACTTAT 1325
      |||
Db      707 eHisThrAlaGluLeuLeuAlaCysPheAlaArgSerArgSerGlyAlaLysLeuIl 727
      |||
QY      1326 CGGAACCGCAACTCTGTTGCTCTCTCGGAATACACTCTCTTCCATTTGGCTGTAGG 1385
      |||
Db      727 eGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLeuLeuGl 747
      |||
QY      1386 GTGTGCTGCGCAACTGATCTCTGCGCGGACGTCCTTTGTCTAGTCCCGCTCGCGCTGAA 1445
      |||
Db      747 yCysAlaAlaAenTrpIleLeuArgGlyThrSerPheValTyrValProSerAlaLeuAs 767
      |||
QY      1446 TCCCGCGGACGACCCGCTCTCGGGGCGGTTTGGGGCTCTACCGTCCCTCTTCACTGCC 1505
      |||

```

```

Db      767 nProIlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHisLeuPr 787
      |||
QY      1506 GTTCGGCGCACCGGCGCACCTCTCTTTACCGGTCTCTCCCGTATGTGCTTCTCTCA 1565
      |||
Db      787 oPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValProSerHi 807
      |||
QY      1566 TCTGCGGACCGTGTGACCTTCGCTTACCTCTGCACTCTGCACTCGACGAGACCCG 1620
      |||
Db      807 sLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPro 825
      |||
RESULT 6
T13468
DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (isolate 09D09HCC)
C:Species: hepatitis B virus, HBV
A:Variety: isolate 09D09HCC
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C:Accession: T13468
R:Takahashi, K.; Akahane, Y.; Hino, K.; Ohta, Y.; Mishiro, S.
Arch. Virol. 143, 2313-2326, 1998
A:Title: Hepatitis B virus genomic sequence in the circulation of hepatocellular carcin
A:Reference number: Z17684; MUID:99129050; PMID:9930189
A:Accession: T13468
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-827 <TAK>
A:Cross-references: UNIPROT:Q9YZU3; UNIPARC:UPI00000F479C; EMBL:AB014368; NID:G3551304;
A:Experimental source: Japanese patient with hepatocellular carcinoma isolate 09D09HCC
C:Genetics:
A:Gene: P
A:Introns: 303/3
C:Superfamily: hepatitis virus DNA-directed DNA polymerase
C:Keywords: DNA biosynthesis; nucleotidyltransferase

Alignment Scores:
Pred. No.:      3 7e-190      Length:      827
Score:          2678.50      Matches:     504
Percent Similarity: 95.91%      Conservative: 12
Best Local Similarity: 93.68%      Mismatches:  18
Query Match:     45.34%      Indels:      5
DB:              2          Gaps:         1

US-10-761-006A-1 (1-3215) x T13468 (1-827)

QY      19  AGCTCTGCTAGATCCCGAGGTGAGGG-----CCTATATTTTCTGCTGGTGGCTC 69
      |||
Db      291 ThrSerLysArgGlnSerSerGlyHisAlaValGluLeuHisProCysTrpTrpLeu 310
      |||
QY      70  GAGTTCGGGAACAGTAACCCCTGTTCCGACTACTGCTCTCCCATATCGTCAATCTTCTC 129
      |||
Db      311 GlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleValAsnLeuLeu 330
      |||
QY      130 GAGGACTGGGGACCTGCAACCAATGGAGAACCAACATCAGGATTCCTAGGACCCCT 189
      |||
Db      331 GluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIleProArgThrPro 350
      |||
QY      190 GCTCGTGTATACAGGGGGGTTTTTCTCGTTTGAAGAATCTCTCACAAATCCCGCAGAGTCT 249
      |||
Db      351 AlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThrThrGluSer 370
      |||
QY      250 AGACTCTG-GTGACTTCTCTCAATTTTCTAGGGGAGGACCAACGCTGTTCTTGGCCAAA 308
      |||
Db      371 ArgLeuValValAspPheSerGlnPheSerArgGlySerThr***Val-SerTrpProLy 390
      |||
QY      309 ATTGCGAGTCCCAACCTCCAATCACTCACCAACCTTGTGTCTCCCAATTTGTCCTGGCT 368
      |||
Db      390 sPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeuSerTrpLe 410
      |||
QY      369 ATCGCTGGAATGTCTGCGGCGGTTTTTATCATATTCCTCTTTCATCTGCTGCTATGCTCA 428
      |||
Db      410 uSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAlaMetProHi 430
      |||
QY      429 TCTTCTTGTGTGGTCTTCTCTGGGACTACCAAGGTATGTTCGGCTTGTCTCTACTTCCAG 488
      |||

```



```
Db 430 sLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSerThrSerAr 450
Qy 489 GAACATCAACACACAGCAGCGGCCATGCAAGACTGCAAGACTCCAGCTCAAGAAACT 548
Db 450 GAsn**AsnTyrGlnHisGlyThrMetGlnAspLeuHisGluSerCysSerArgAsnLe 470
Qy 549 CTAGCTTTCCCTCTGTTGCTGTACAAAACCTTCGGACGGAACCTGACACTGTATTCCTCA 608
Db 470 uTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLeuTyrSerHi 490
Qy 609 TCCCATCATCTCGGCTTTTCGCAAGATTCTCTATGGAGTGGGCTCAGTCCGTTTCCTCT 668
Db 490 sProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerProPheLeuLe 510
Qy 669 GGCTCAGTTACTAGTCCATTTCTTCAGTGGTTGCTAGGCTTTCCCACTGTTTGGC 728
Db 510 uAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHisCysLeuAl 530
Qy 729 TTTTCAGTTATATGATGATGTTGTTGGGGCGAAGTCTGTACAAATCTTCAGTCCCT 788
Db 530 aPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLeuGluSerLe 550
Qy 789 TTTTACCTCTATTACCAATTTCTTTGTCTTTGGGTATACATTTAAACCTTAATAAAC 848
Db 550 uPheThrAlaIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnProHisLysTh 570
Qy 849 CAACAGTTGGGGCTACTCCCTTAATCTTATGGGATATGTAATCGAAAGTTGGGTACTTT 908
Db 570 xLysArgTyrGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTyrGlyThrLe 590
Qy 909 ACCGAGACATATTGTTACTTAAACCTCAAGCAATGTTTCGAAACCTGCTGTAAATAG 968
Db 590 uProGlnGluHisIleValGlnLysIleLysGlnCysPheArgLysLeuProValAsnAr 610
Qy 969 ACCTATTGATTGGAAGTATGCTCAAGAAATTTGCGGCTTTTGGGCTTTGTCGCCCTTT 1028
Db 610 sProIleAspTyrLysValCysGln**IleValGlyLeuLeuGlyPheAlaAlaProPh 630
Qy 1029 TACAATATGGTCTATCTGCTTATGCTTATATGATGATATCAATCTTAAGCAGGC 1088
Db 630 eThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnAlaLysGlnAl 650
Qy 1089 TTTTCACTTTCTGGCAACTACAGGCTTCTGTGTAAACAATATCTGAACCTTTACCC 1148
Db 650 aPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuHisLeuTyrPr 670
Qy 1149 GTTTCGCCGCAACGGTCCGCTCTGCAAGTGTGCTGACGCAACCCCACTGGATG 1208
Db 670 oValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrProThrGlyTr 690
Qy 1209 GGGCTTGGCCATAGGCCATCAGCGATGGCTGGAAACCTTTCTGGCTCTCTCGCGATCCA 1268
Db 690 pGlyLeuAlaIleGlyHisArgMetArgGlyThrPheValAlaProLeuProIleHi 710
Qy 1269 TACTCGGCAACTCTAGCAGCTTGTGCTCGCAGCGGCTCGAGCAAACTATCGG 1328
Db 710 sThrAlaGluLeuAlaIaCysPheAlaArgSerArgSerGlyAlaLysLeuIleGl 730
Qy 1329 AACCGACAACCTGTGTGCTCTCTCGGAATACACCTCTCTTCATGCTGTAGGGTG 1388
Db 730 yThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProThrLeuLeuGlyCy 750
Qy 1389 TGCTGCCAACTCGGATCTCGCGGACGTCCTTTGTCTACGTCCTCGCGCTGGAATCC 1448
Db 750 sAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAlaLeuAsnPr 770
Qy 1449 CGCGGACGACCGCTCTCGGGCCGTTTGGGGCTCTACCGTCCCTCTTCTCATCTGCCGTT 1508
Db 770 oAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHisLeuProPh 790
Qy 1509 CGGCGGACACGCGGGCCACTCTCTTACGCGGTCTCCCGGTATGCGCTTCTCATCT 1568
Db 790 eArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValProSerHisLe 810
```

```
Qy 1569 GCCGGACCGTGTGCACCTTCGCTTACCTCTGCAAGTCGCGATGAGAGACCACCG 1620
Db 810 uProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPro 827
```

RESULT 7

```
S71785
DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype ayw, isolate pati
C;Species: hepatitis B virus, HBV
A;Variety: subtype ayw, isolate patient C1005
C;Date: 06-Dec-1996 #sequence_revision 06-Dec-1996 #text_change 09-Jul-2004
C;Accession: S71785
R;Preisler-Adams, S.; Schlaver, M.J.; Peters, T.; Hettler, F.; Rasenack, J.
submitted to the EMBL Data Library, March 1993
A;Description: Identification and sequence analysis of hepatitis B virus DNA in immunol
A;Reference number: S32202
A;Accession: S71785
A;Molecule type: DNA
A;Residues: 1-832 <PRE>
A;Cross-references: UNIPROT:Q8JN11; UNIPROT:Q9IF40; UNIPROT:Q96846; UNIPROT:Q67892; UNI
PROT:Q9QAG0; UNIPROT:Q9QAF3; UNIPROT:Q8QXQ1; UNIPROT:O11885; UNIPROT:O56654; UNIPROT:O5
A;Experimental source: subtype ayw, isolate patient C1005
C;Superfamily: hepatitis virus DNA-directed DNA polymerase
C;Keywords: DNA biosynthesis; nucleotidyltransferase
```

Alignment Scores:

Pred. No.:	2,42e-188	Length:	832
Score:	2654.00	Matches:	492
Percent Similarity:	95.56%	Conservative:	25
Best Local Similarity:	90.94%	Mismatches:	23
Query Match:	44.93%	Indels:	2
DB:	2	Gaps:	0

US-10-761-006A-1 (1-3215) x S71785 (1-832)

```
Qy 1 CTCACAACTTCCACCAAGCTCTGTAGATCCAGGGTGGAGGCGCTATATTTTCCTGC 60
Db 293 LeuHisAsnPheProAsnSerAlaArgSerGlnGlyGluArgProAlaPheProCys 312
Qy 61 TGTGGCTCCAGTTCGCGAACAGTAAACCTGTTCGAGTACTGCTCTCCCATATCGTC 120
Db 313 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerHisIleVal 332
Qy 121 AATCTTCTCAGAGCTGGGGACCTCGACCGAACAATCGAGAACACAACATCAGGATTCCT 180
Db 333 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisIleArgIlePro 352
Qy 181 AGGACCCCTGCTGTTTACAGGCGGGTTCCTCGTTCACAAGAATCTCTACAATACC 240
Db 353 ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 372
Qy 241 GCAGAGTCTAGACTCTG-GTGGACTTCTCTCAATTTTCTAGGGGGAGCACCCACGTGTC 299
Db 373 AlaGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnHisArgVal-Se 392
Qy 300 CTGGCCAAATTCGAGTCCCAACCTCCAATCATCATCACCACCTCTGCTCTCCAATTT 359
Db 392 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 412
Qy 360 GTCTGGCTATCGCTGGATGTCGTGGCGGTTTTCATATTCCTCTCTCATCTGCTGTC 419
Db 412 uSerTrpLeuSerLeuAspValSerAlaIaPheTyrHisLeuProLeuHisProAlaSe 432
Qy 420 TATGCTCTCATCTCTTCTGTTGCTTCTCTGAGCTACCAAGGTATGTTGCCGTTTGTCTC 479
Db 432 rMetProHisLeuLeuValGlySerThrGlyLeuSerArgTyrValAlaArgValSerSe 452
Qy 480 TACTTCCAGGAACATCAACCAACGACGCGGGCCATGCAAGACTGACGAGCTCTGCTC 539
Db 452 rAsnSerArgIlePheAsnHisGlnArgGlyThrMetGlnAsnLeuHisAspTyrCysSe 472
Qy 540 AAGGAAACTCTACGTTTCCCTCTGTTGCTGTACAAAACCTTCGGACGGAACTGCACTT 599
```

Db 472 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrGlnThrPheGlyArgGlyLeuHisLe 492

QY 600 GTATTCCCATCCCATCTCTGGGCTTTGCGAAGATTCCTATGGGAGTGGGCTCAGTCC 659

Db 492 uTyrSerHisProIleLeuGlyPheArgIysIleProMetGlyValGlyLeuSerPr 512

QY 660 GTTTCCTCCGGCTCAGTTACTAGTGCCTATTTGTTTCAGTGGTTCGTAGGGCTTCCCA 719

Db 512 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 532

QY 720 CTCTTTGGCTTTTCAGTTATATGATCATCTGGTATTGGGGCGCAAGCTCTGTACAACTCT 779

Db 532 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaIysSerValGlnHisLe 552

QY 780 TGAGTCCCTTTTACCTCTATACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCTC 839

Db 552 uGluSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 572

QY 840 TAATAAACCNAAGTTGGGCTACTCCCTTAACCTTCATGGGATATGTAATTGGAAGTTG 899

Db 572 oAsnLeuThrLysArgTyrGlyTyrSerLeuHisPheMetGlyTyrValIleGlySerTy 592

QY 900 GGGTACTTTTACCCGACCAATATGTTACTTAAACTCAAGCAATGTTTTCGAAACCTGCC 959

Db 592 rGlySerLeuProGlnAspHisIleIleGlnIysIleIysGluCysPheArgIysLeuPr 612

QY 960 TGTAAATAGACCTATTGATTGGAAGTATGTCAAAGAATTGTGGGTCTTTTGGGCTTTGC 1019

Db 612 oIleAsnArgProIleAspTrpIysValCysGlnArgIleValGlyLeuLeuGlyPheAl 632

QY 1020 TGCCCTTTTACCAATGTGGCTATCTCTGCTTGATGCTTTATATGATGATATCAATC 1079

Db 632 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 652

QY 1080 TAAGCAGGCTTTTCACTTTCTCGCAACTTACAGGCTTTCTGTGTAAACAATATCTGAA 1139

Db 652 rLysGlnAlaPheThrPheSerProThrTyrIysAlaPheLeuCysIysGlnTyrLeuAs 672

QY 1140 CCTTTACCCCGTTGCCGCAACGGTCCGGTCTCTGCCAAGTGTTTGTGACGCAACCCC 1199

Db 672 nLeuTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 692

QY 1200 CACTGGATGGGCTTGGCCATAGGCCATCAGCCATGCGCTGGAACTTTCTGGCTCTCT 1259

Db 692 oThrGlyTyrGlyLeuValMetGlyHisGlnArgMetArgGlyThrPheLeuAlaProle 712

QY 1260 GCGATCCATACCTGCGGAATCTCTAGCAGCTGTTTGTCTCGCAGCCGGTCTGGAGCAA 1319

Db 712 uProIleHisThrAlaGluLeuLeuAlaIaCysPheAlaArgSerArgSerGlyAlaAs 732

QY 1320 ACTTATCGGAACCGACAACCTGTGTCTCTCTCGGAATACACCTCTCTTCCATGGCT 1379

Db 732 nIleLeuGlyThrAspAsnSerValValLeuSerArgIysTyrThrSerPheProTrple 752

QY 1380 GCTAGGGTGTGTGCCAATCTGATCTCTGGCGGAGCTCTTTGTCTAGTCCCGTCGGC 1439

Db 752 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 772

QY 1440 GCTGAATCCCGGAGCAGCCCGTCTCGGGCGCTTTGGGCTCTACCGTCCCTCTTCCA 1499

Db 772 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyIlePheArgProLeuLeuAr 792

QY 1500 TCTGCGGTTCCGCGCACCAACCGGCGCACCTCTCTTTACGGGGTCTCCCGGTATGTGCC 1559

Db 792 gLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaAspSerProSerValPr 812

QY 1560 TTCTCATCTCCGAGCCGTGTGCACTTCGCTTACCTCTGACGTCGATGGAGACCACC 1619

Db 812 oSerHisLeuProValArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 832

QY 1620 G 1620

Db 832 o 832

RESULT 8

JDVLVB

DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype ayw, strain pH320)
C:Species: hepatitis B virus, HBV
A>Note: host Homo sapiens (man)
C>Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 09-Jul-2004
C:Accession: A00703
R:Richko, V.; Pushko, P.; Dreilina, D.; Pumpen, P.; Gren, E.
FEBS Lett. 185, 208-212, 1985
A:Title: Subtype ayw variant of hepatitis B virus: DNA primary structure analysis.
A:Reference number: A05237; MUID:85204397; PMID:3996597
A:Accession: A00703
A:Molecule type: DNA
A:Residues: 1-832 <BIC>
A:Cross-references: UNIPROT:P03156; UNIPARC:UPI00001710E8; GB:X02496; NID:G62280; PIDN:C
C:Superfamily: hepatitis virus DNA-directed DNA polymerase
C:Keywords: DNA biosynthesis; nucleotidyltransferase

Alignment Scores:

Pred. No.: 3,4e-188 Length: 832
Score: 2652.00 Matches: 494
Percent Similarity: 95.38% Conservative: 22
Best Local Similarity: 91.31% Mismatches: 24
Query Match: 44.90% Indels: 2
DB: 1 Gaps: 0

US-10-761-006A-1 (1-3215) x JDVLVB (1-832)

QY 1 CTCCGACACATTCACCAAGCTCTGTAGATCCCGGGGTGAGGGCCCTATATTTCCTGC 60

Db 293 LeuHisAsnLeuProProAsnSerAlaArgSerGlnSerGluArgProValPheProCys 312

QY 61 TGGTGGCTCCAGTTCGCGAACAGTAAACCCCTGTTCCGACTACTGCTCTCCCATATCTGC 120

Db 313 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerHisIleVal 332

QY 121 AATCTTCTCGAGGACTGGGGACCCCTGCAACCAATGGAGAACACCAATCAGGATTCCT 180

Db 333 AsnLeuLeuGluAspTrpGlyProCysAlaGluHisGlyGluHisIleArgIlePro 352

QY 181 AGGACCCCTGCTCGTGTACAGGGGGGTTTTCTCGTTGACAGAAATCCTCACAAATACC 240

Db 353 ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 372

QY 241 GCAGAGTCTAGACTCTG-GTGAGCTTCTCAATTTCTAGGGGAGCACCCACGCTGTC 299

Db 373 AlaGluSerArgLeuValValAspPheSerGlnPheSerArgIysAsnTyrArgVal-Se 392

QY 300 CTGGCCAAAATTGCGAGTCCCCAACCTCCAATCACTCAACCACTCTTGTCTCCAAATTT 359

Db 392 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 412

QY 360 GTCTGGCTATCGCTGGATGTGTCTGCGCGGTTTTATCATATTCCTCTCTCTGCTGC 419

Db 412 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaAl 432

QY 420 TATGCTCATCTCTTGTGTGTCTCTGGAATACCAAGGTATGTTCGCCGTTTGTCTCTC 479

Db 432 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaArgLeuSerSe 452

QY 480 TACTTCCAGGAACATCAACACCACCGGCGCCATGCAAGACCTGCAAGCTCTCTGCTC 539

Db 452 rAsnSerArgIlePheAsnTyrGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe 472

QY 540 AAGGAACTCTACGTTTCCCTCTTGTCTGTACAAACCTTCGGACCGGAACCTGCACTT 599

Db 472 rArgAsnLeuTyrValSerLeuLeuLeuTyrGlnThrPheGlyArgLysLeuHisLe 492

QY 600 GTATTCCCATCCCATCTCTGGGCTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCC 659

Db 492 uTyrSerHisProIleIleLeuGlyPheArgIysIleProMetGlyValGlyLeuSerPr 512

```
QY 660 GTTCTCTCGCTCAGTTTACTAGTCCATTGTTCTCAGTGGTTCGTAGGCGTTTCCGCCA 719
Db 512 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProH 532
QY 720 CTGTTTGGCTTTCAGTTATATGATGATGATGATGATGATGATGATGATGATGATGAT 779
Db 532 sCysLeuAlaPheSerThrMetAspAspValValLeuGlyAlaIleValGlnHisLe 552
QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTGTTCTTTGGGTATACATTAACCC 839
Db 552 uGluSerLeuPheThrAlaValThrAsnLeuLeuSerLeuGlyIleHisLeuAsnPr 572
QY 840 TAATAAACCAACGTTGGGCTACTCCCTTAACCTTCATGGGATATGTTAATTCGAAGTTG 899
Db 572 oAsnLysThrLysArgTrpGlyThrSerLeuHisPheMetGlyThrValIleGlyCysTy 592
QY 900 GGGTACTTACCGCAGGAACATATTGTAATAAATCAAGCAATGTTTTCGAAACCTGCC 959
Db 592 rGlySerLeuProGlnAspHisIleIleGlnLysIleGlyCysPheArgLysLeuPr 612
QY 960 TGTAAATPAGACCTATTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1019
Db 612 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 632
QY 1020 TGGCCCTTTTACACAAATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1079
Db 632 aAlaProPheThrGlnCysGlyThrProAlaLeuMetProLeuThrAlaCysIleGlnSe 652
QY 1080 TAAGCAGGCTTTCATCTTCTCGCAACTTACAGGCTTCTGTTGTAACAAATATCTGAA 1139
Db 652 rLysGlnAlaPheThrPheSerProThrLysAlaPheLeuCysLysGlnThrLeuAs 672
QY 1140 CCTTTACCCCTTTCGCCGCAACCGTCCGCTCTCTGCAAGTGTGCTGACGCAACCCC 1199
Db 672 nLeuThrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 692
QY 1200 CACTGGATGGGCTTGCCATAGCCATCAGCCATGCTGGAACCTTCTGGCTCTCTCT 1259
Db 692 oThrGlyTrpGlyLeuValMetGlyHisGlnArgMetArgGlyThrPheLeuAlaArgLe 712
QY 1260 GCCGATCCATCTGCGCAACTCTAGCAGCTTGTGTTGCTCGCAGCGGTCTCGAGCAA 1319
Db 712 uProIleHisThrAlaGlnLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaAs 732
QY 1320 ACTTATCGAACCGCAACTCTGTGTCTCTCTCGGAATATACACCTCTTTCATGGCT 1379
Db 732 nIleLeuGlyThrAspAsnSerValValLeuSerArgLysThrThrSerThrProThrLe 752
QY 1380 GCTAGGTGTGCTGCCAACTGGATCTCGCGGAGCTCTCTTGTCTACGTCGCGTCCG 1439
Db 752 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValThrProSerAl 772
QY 1440 GCTGAATCCCGCAGCAGCAGCTCTCGGCGCGTTTGGGGCTCTACCGTCCCTCTTCA 1499
Db 772 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuSerArgProLeuLeuAr 792
QY 1500 TCTGCCGTTCCGGCGCAGCAGCGGCGCACCTCTCTTTACGCGGTCTCCCGGTATGSCC 1559
Db 792 gLeuProPheArgProThrThrGlyArgThrSerLeuThrAlaAspSerProSerValPr 812
QY 1560 TTCTCATCTGCGGACCGTGTGCACTTCTGCTTCACTCTGACGTCGATCGGAGACACC 1619
Db 812 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 832
QY 1620 G 1620
Db 832 o 832
```

RESULT 9
S47406

DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype ayw4)

C;Species: hepatitis B virus, HBV

A;Variety: subtype ayw4

C;Date: 23-Nov-1994 #sequence_revision 22-Nov-1996 #text_change 09-Jul-2004
C;Accession: S47406
R;Pluclenniczak, A.
submitted to the EMBL Data Library, August 1994
A;Description: Molecular cloning and sequencing of two complete genomes of polish isola
A;Reference number: S47404
A;Accession: S47406
A;Molecule type: DNA
A;Residues: 1-832 <PLU>
A;Cross-references: UNIPROT:Q67892; UNIPARC:UPI00000BF97B; EMBL:235716; NID:9527435; PI
A;Experimental source: subtype ayw4
C;Genetics:
A;Gene: P
C;Superfamily: hepatitis virus DNA-directed DNA polymerase
C;Keywords: DNA biosynthesis; nucleotidyltransferase

Alignment Scores: 4.03e-188 Length: 832
Pred. No.: 2651.00 Matches: 494
Score: 2651.00
Percent Similarity: 95.19% Conservative: 21
Best Local Similarity: 91.31% Mismatches: 25
Query Match: 44.88% Indels: 2
DB: 1 Gaps: 0

US-10-761-006a-1 (1-3215) x S47406 (1-832)

```
QY 1 CTCACAACATTCACCAAGCTCTGTAGATCCAGGTGAGGGGCTATATTTTCTCTGC 60
Db 293 LeuHisAsnLeuProAsnSerSerArgSerGlnGlyGluArgProValPheProCys 312
QY 61 TGTGTGCTCCAGTTCGGACAGTAACCTGTTTCGACTACTGCTCTCCATATGCTC 120
Db 313 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspThrCysLeuSerHisIleVal 332
QY 121 AATCTTCTCGAGACTGGGGACCTCGACCGAATCGGAGAACCAACATCAGGATTCCT 180
Db 333 AsnLeuLeuGluAspTrpGlyProCysAlaGluHisGlyGluHisIleArgIlePro 352
QY 181 AGGACCCCTGCTCGTGTTCAGCGGGGTTTCTCTCAATTTCTAGGGGGAGCACCAGTGCTC 240
Db 353 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 372
QY 241 GCAGACTAGACTCTG- GTGGACTTCTCTCAATTTCTAGGGGGAGCACCAGTGCTC 299
Db 373 AlaGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnThrArgVal-Se 392
QY 300 CTGGCCAAATTCGACAGTCCCAACTCCAACTCACTCACCAACTCTGTGCTCCCAATT 359
Db 392 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuSerSerAsnLe 412
QY 360 GTCCTGGCTATCGCTGGATGTGTCTGCGCGTTTATCATATTTCTCTTCTCTCTGCTGC 419
Db 412 uSerTrpLeuSerLeuAspValSerAlaAlaPheThrHisLeuProLeuHisProAlaAl 432
QY 420 TATGCTCATCTTCTGTTGGTCTTCTGACTACCAAGGTATGTTCCCGTGTGCTCTC 479
Db 432 aMetProHisLeuLeuValGlySerGlyLeuSerArgThrValAlaArgLeuSerSe 452
QY 480 TACTTCCAGGAACATCAACCAACGACGCGGGCCATGCAAGACTCGCACACTCTCTGCTC 539
Db 452 rAsnSerArgIlePheAsnAsnGlnArgGlyThrMetGlnAsnLeuHisAspThrCysSe 472
QY 540 AAGGAAACTCTACGTTTCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 599
Db 472 rArgAsnLeuThrValSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 492
QY 600 GTATTCCCATCCCATCATCTCTGGGCTTCCAGATTCCTATCGGAGTGGGCTTCAGTCC 659
Db 492 uThrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 512
QY 660 GTTCTCTCGCTCAGTTTACTAGTGCATTTGTTTCTAGTGGTTCGTAGGCGTTTCCGCCA 719
Db 512 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProH 532
```



```
Db 532 sCysLeuAlaPheSerThrMetAspValLeuGlyAlaLysThrValHisLeu 552
QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTGTTGGGTATACATTTAAACCC 839
Db 552 uGluSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 572
QY 840 TAATAAACCAGAGTTGGGCTACTCCCTTAACCTTCATGGGATATGTAATTCGAAGTTG 899
Db 572 oAsnLysThrLysArgTrpGlyThrSerLeuHisPheMetGlyThrValIleGlyCysTy 592
QY 900 GGGTACTTTACCGCAGCAATATGTAATAAATCAAGCAATGTTTTCGAAACCTGCC 959
Db 592 rGlySerLeuProGlnAsnHisIleIleGlnLysIleLysGluCysPheArgLysLeuPr 612
QY 960 TGTAAATAGACCTATTGATTTGGAAGPATGTCAAGAAATTTGTGGCTTTTGGCTTTGC 1019
Db 612 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 632
QY 1020 TGCCCTTTTACCAATGTGGCTATCTCGCTTGATGCTTATATGCAATGTATACATC 1079
Db 632 aAlaProPheThrGlnCysGlyThrProAlaLeuMetProLeuThrAlaCysIleGlnSe 652
QY 1080 TAAGCAGCTTTTCACTTTCTCGCAACTTACAAGCCCTTTCTGTGTAAACAATATCTGAA 1139
Db 652 rLysGlnAlaPheThrPheSerProThrThrLysAlaPheLeuCysLysGlnThrLeuAs 672
QY 1140 CTTTATACCGTTGCGCGCAACGGTCCGGTCTCTGCAAGTGTGTTGCTGACCAACCCC 1199
Db 672 nLeuThrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 692
QY 1200 CACTGGATGGGCTTGCCATAGGCCATCAGCGCATCGCTGGAACCTTTCTGGCTCTCT 1259
Db 692 oThrGlyTrpGlyLeuValMetGlyHisGlnArgMetArgGlyThrPheGlnAlaProLe 712
QY 1260 GCGATCCATACTCGCGAATCTCTAGCAGCTTTGTTGCTCGCAGCGGCTCGAGCAAA 1319
Db 712 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaAs 732
QY 1320 ACTTATCGGAACCGCAACTCTGTTGCTCTCTCGGAATATACACTCTCTTCCATGGCT 1379
Db 732 nIleLeuGlyThrAspAsnSerValValLeuSerArgLysThrThrSerPheProTrpLe 752
QY 1380 GTTAGGTGTGCTGCCAATCGATCTCTGCGGGAGCTCTTTGCTACGTCCGCTCGGC 1439
Db 752 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValThrValProSerAl 772
QY 1440 GCTGAATCCCGCGCAGCAGCTCTCGGGCGGTTTGGGGCTCTACCGTCCCTCTCTCA 1499
Db 772 aLeuAsnProAlaAspAspProSerArgLysArgLeuGlyLeuSerArgProLeuLeuAr 792
QY 1500 TCTCCGTTCCGGCGCAGCAGCGGCGCACCTCTCTTTACGCGTCTCCCGATGTGCC 1559
Db 792 gLeuProPheArgProThrThrGlyArgThrSerLeuThrAlaAspSerProSerValPr 812
QY 1560 TTCTCATCTCGCGACCGTGTGCACTTCCTGCTTCACTCTGCACTCGCATCGGATGAGCACCC 1619
Db 812 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 832
QY 1620 G 1620
Db 832 o 832
```

RESULT 13

S20757

DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype ayw, patient E)

C:Species: hepatitis B virus, HBV

A:Variety: subtype ayw, patient E

C:Date: 20-Feb-1995 #sequence_revision 22-Nov-1996 #text_change 09-Jul-2004

C:Accession: S20757

R:Lai, M.E.; Mazzoleni, A.P.; Balestrieri, A.; Mellis, A.; Porru, A.

submitted to the EMBL Data Library, March 1992

A:Description: Sequence analysis of HBV genomes isolated from patients with HBSAg negati

A:Reference number: S20745
A:Accession: S20757
A:Molecule type: DNA
A:Residues: 1-832 <LAI>
A:Cross-references: UNIPROT:Q67882; UNIPARC:UPI00000F25B1; EMBL:X65259; NID:959439; PID
A:Experimental source: subtype ayw, patient E
C:Genetics:
A:Gene: P
C:Superfamily: hepatitis virus DNA-directed DNA polymerase
C:Keywords: DNA biosynthesis; nucleotidyltransferase

Alignment Scores:
Pred. No.: 4,38e-187 Length: 832
Score: 2637.00 Matches: 494
Percent Similarity: 95.19% Conservative: 21
Best Local Similarity: 91.31% Mismatches: 25
Query Match: 44.64% Indels: 2
DB: 1 Gaps: 0

US-10-761-006A-1 (1-3215) x S20757 (1-832)

```
QY 1 CTCACAAATTCACCAAGCTCTGTAGATCCAGGGTGAGGGGCTATATTTTCTCTGC 60
Db 293 LeuHisAsnLeuProAsnSerAlaArgSerGlnSerGluArgProValPheProCys 312
QY 61 TGTGGCTCCAGTTCGGGAACAGTAAACCTGTTCGACTACTCTCCATATCGTC 120
Db 313 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerLeuIleVal 332
QY 121 AATCTTCTCGAGACTGGGGACCTCCACGAACTGGAGAACACAAACATCAGGATTCCT 180
Db 333 AsnLeuArgGluAspTrpGlyProCysThrGluHisGlyGluHisIleArgIlePro 352
QY 181 AGGACCTCTCGTGTGTACAGCGGGGTTTTCTCTGTGACAAAGAAATCTCAATACC 240
Db 353 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 372
QY 241 GCGAGTCTAGACTCTG-GTGGACTTCTCTCAATTTCTAGGGGGAGCACCACGCTTC 299
Db 373 AlaGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnTyrArgVal-Se 392
QY 300 CTGGCCAAATTCGCGAGTCCCAACTCCAACTCACTCACCAACTCTTTGCTCCCAATT 359
Db 392 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 412
QY 360 GTCTCTGCTATCTCGTGTGTGTCTCGCGGTTTTATCATATTTCTCTTCTCATCTGCTC 419
Db 412 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaAl 432
QY 420 TATGCTCATCTCTTCTTGTGGTCTTCTTGAGCTACCAAGTATGTTGCCCGTTGCTCTC 479
Db 432 aMetProHisLeuLeuValGlySerGlyLeuSerArgTyrValAlaArgLeuSerSe 452
QY 480 TACTTTCAGGAACATCAACACACAGCAGCGGGCCATGCAAGACTGCACCACTCTCTGCTC 539
Db 452 rAsnSerArgLysPheAsnAsnGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe 472
QY 540 AAGGAAACTCTAGTTTCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 599
Db 472 rArgAsnLeuTyr-ValSerLeuLeuLeuLeuTyrGlnThrPheGlyArgLysLeuHisLe 492
QY 600 GTATTCCTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
Db 492 uTyrSerHisProIleLeuLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 512
QY 660 GTTTCCTCTGGCTCAGTTTACTAGTGCATTTGTTCAGTGTGTTCTGAGGCTTTCCCCCA 719
Db 512 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHI 532
QY 720 CTGTTTGGCTTTTCTAGTTATATGATGATGATGATGATGATGATGATGATGATGATGAT 779
Db 532 sCysLeuAlaPheSerThrMetAspValValLeuGlyAlaLysSerValGlnHisLe 552
```

780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCC 839
 552 uGluSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 572
 840 TAAATAAACCAACGTTGGGCTACTCCCTTAACCTCATGGGATATGTAATTTGGAAGTTG 899
 572 oAenLySthrLySArgTrpGlyTyrSerLeuAsnPheMetGlyTyrIleIleGlySerPr 592
 900 GCGTACTTTACCGCAGCAACATATTGTACTAAACCTCAAGCAATGTTTTCGAAAACTGCC 959
 592 pGlyThrLeuProGlnAspHisIleValGlnLySleIleLySGLuCySPheArgLySLeuPr 612
 960 TGTAAATAGACCTATTGATTGGAAGTATGTCAAGAATTTGGGGTCTTTTGGGCTTTCG 1019
 612 oValAsnArgProIleAspTrpLySValTrpGlnArgIleValGlyLeuLeuGlyPheAl 632
 1020 TGCCCTTTTACCAATGGCTATCTCTGCTGTAGCTTTATATGATGATATCAATC 1079
 632 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnAl 652
 1080 TAAGCAGGCTTTTACCTTTCTCGCAACTTACAGGCTTTCTGTGTAAACAATATCTGAA 1139
 652 aLySGLnAlaPheThrPheSerProThrTyrLySAlaPheLeuSerLySGlnTyrWetAs 672
 1140 CCTTTACCCCGTTGCGCGCAACGGTCCGGTCTCTGCCAAGTGTGTGCTGACGCAACCC 1199
 672 nLeuTyrProValAlaArgGlnArgProGlyLeuCySGLnValPheAlaAspAlaThrPr 692
 1200 CACTGATGGGCTTGGCCATAGGCCATCAGCGCATGGCTGGAACTTTCTGGCTCTCT 1259
 692 oThrGlyTrpGlyLeuAlaIleGlyAsnGlnArgMetArgGlyThrIleValAlaProLe 712
 1260 GCGGATCCCATCTGCGCAACTCTCTAGCAGCTTTGTTGCTCGCAGCGGTCTGGAGCAA 1319
 712 uProIleHisThrAlaGluLeuAlaIleGlyAsnGlnArgMetArgGlyThrIleValAla 732
 1320 ACTTATCGGAACCGCAACTCTGTGTCTCTCTCGGAATATACACCTCTCTTCCATGGCT 1379
 732 sLeuIleGlyThrAspAsnSerValValLeuSerArgLySThrThrSerPheProTrpLe 752
 1380 GCTAGGGTGTGTCGCAACTGATCTGCGCGGAGCTCTTTGCTAGCTCGCTCGCGG 1439
 752 uLeuGlyCySthrAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 772
 1440 GCTGAATCCCGCGAGCAGCCGCTCGCGGCGGCTTTGGGCTCTACCGTCCCTCTTCCA 1499
 772 aLeuAsnProAlaAspProSerArgLySArgLeuGlyLeuSerArgProLeuLeuAr 792
 1500 TCTGCCGTTTCCGCGCAGCAGCGGCGCACCTCTCTTTACGGGCTCTCCCGGTATGGCC 1559
 792 gLeuProPheGlnProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 812
 1560 TTCTCATCTGCGGACCGTGTGACCTTCTGCTTACCTCTGACGTCGATGGAGCACCC 1619
 812 oSerHisLeuProValArgValHisPheAlaSerProLeuHisIleAlaTrpArgProPr 832
 1620 G 1620
 832 o 832

RESULT 14
 S67505
 DNA-directed DNA polymerase (BC 2.7.7.7) - hepatitis B virus (subtype ayw3, isolate Hope
 C:Species: hepatitis B virus, HBV
 A:Variety: subtype ayw3, isolate Hope CH1357
 C:Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 09-Jul-2004
 C:Accession: S67505
 R:Norde, H.; Ebert, J.W.; Fields, H.A.; Mushahwar, I.K.; Magnus, L.O.
 Virolgy 218, 214-223, 1996
 A:Title: Complete sequencing of a gibbon hepatitis B virus genome reveals a unique genob
 A:Reference number: S67503; MUID:96207410; PMID:8615024
 A:Accession: S67505
 A:Molecule type: DNA

A:Residues: 1-832 <NOR>
 A:Cross-references: UNIPROT:P87744; UNIPARC:UPI0000074B8; EMBL:U46935; NID:g1814218; PI
 A:Experimental source: subtype ayw3, isolate Hope CH1357
 A:Note: the authors translated the codon GAA for residue 290 as Ala
 C:Genetics:
 A:Gene: P
 A:Introns: 292/3
 C:Superfamily: hepatitis virus DNA-directed DNA polymerase
 C:Keywords: DNA biosynthesis; nucleotidyltransferase

Alignment Scores:
 Pred. No.: 2,03e-186 Length: 832
 Score: 2628.00 Matches: 490
 Percent Similarity: 95.19% Conservative: 25
 Best Local Similarity: 90.57% Mismatches: 25
 Query Match: 44.49% Indels: 2
 DB: 2 Gaps: 0

US-10-761-006A-1 (1-3215) x S67505 (1-832)

Qy 1 CTCCACAACATTCCACCAAGCTCTCTAGATCCAGGGTGAGGGCGCTATATTTCTCTGC 60
 Db 293 LeuTyrSerIleProProAsnSerAlaArgSerGlnSerThrGlyProIleLeuSerCys 312
 Qy 61 TGTGTGGCTCCAGTTCCGGAACAGTAAACCCCTGTTCGACTACTGCTCTCTCCATATCGTC 120
 Db 313 TrpTrpLeuGlnPheArgAsnSerGluProCySAspTyr-CysLeuSerHisLeuVal 332
 Qy 121 AATCTTCTCGAGGACTGGGGACCCCTGCACCGAACAATGAGAGACACACATCAGGATTCCT 180
 Db 333 AsnLeuLeuGluAspTrpGlyProCySThrGluHisGlyGluHisIleArgIlePro 352
 Qy 181 AGGACCCCTGCTCGTGTACAGCGGGGTTTTCTCGTTGACAAGAATCCTCACAAATACC 240
 Db 353 ArgThrProAlaArgValThrGlyValPheLeuValAspLySAsnProHisAsnThr 372
 Qy 241 CGAGAGTCTAGACTCTG-GTGACTTCTCTCAATTTCTTAGGGGAGCACCCACGCTGTC 299
 Db 373 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrArgVal-Se 392
 Qy 300 CTGGGCAAAATTCGAGTCCCCCAACTCCAATCCTCACCACCTCTTGTCTCTCCAAATTT 359
 Db 392 rTrpProLySPhelAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 412
 Qy 360 GTCTTGGCTATCGCTGGATGTCTGCGCGGTTTTATCATATTCTCTTCTATCTCTCTGC 419
 Db 412 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaAl 432
 Qy 420 TATGCTCATCTTCTTGTGTGGTTCTCTGAGACTACCAAGGTATGTGCCCGTTTGTCTTC 479
 Db 432 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaArgLeuSerSe 452
 Qy 480 TACTTTCAGGAACATCAACACGAGCGGGCCATGCAAGACCTGCAGACTCTCTCTGCTC 539
 Db 452 rThrSerArgIleIleAspHisGlnHisGlyThrMetGlnAsnLeuHisAspHisCysSe 472
 Qy 540 AAGGAACTCTACGTTTCCCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 599
 Db 472 rArgAsnLeuPheValSerLeuMetLeuLeuTyrThrPheGlyArgLySLeuHisLe 492
 Qy 600 GTATTTCCTCATCTCTGGGCTTCTGCAAGATTCCTATGGGAGTGGGGCTCAGTCC 659
 Db 492 uTyrSerHisProIleValLeuGlyPheArgLySLeuProMetGlyValGlyLeuSerPr 512
 Qy 660 GTTCTCTCTGGCTCAGTTTACTAGTCCCAATTTGTCAGTGGTTCGTAGGGCTTCCCCCA 719
 Db 512 oPheLeuLeuAlaGlnPheThrSerSerIleCysSerValValArgArgAlaPheProHi 532
 Qy 720 CTGTTTGGCTTTCAGTTATATGATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 779
 Db 532 sCysLeuAlaPheSerTyrWetAspAspLeuValLeuGlyAlaLySerValGlnHisLe 552
 Qy 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTTGGGTATACATTTAAACCC 839

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 27, 2005, 20:40:32 ; Search time 180.238 Seconds
(without alignments)
3320.243 Million cell updates/sec

Title: US-10-761-006A-1_COPY_155_835
Perfect score: 1278
Sequence: 1 ATGGAGACACACATCAGG.....GTCCTTGGGTATACATTAA 681

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2443163 seqs, 439378781 residues
Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DRV=xlp
-Q/cgn2.1/USPTO.spool.p/US10761006/runat 27122005 192900 15743/app query.fasta_1.1102
-DB-A Geneseq -QFWT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPT=0 -LCOPT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10761006 @CGN 1.1 476 @runat 27122005 192900 15743 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A Geneseq 21.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*
9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1230	96.2	400	3	AAY54045 Amino aci
2	1206	94.4	226	1	AAP40310 Hepatitis
3	1206	94.4	226	1	AAB35768 Hepatitis
4	1206	94.4	226	4	AAB98043 Synthetic
5	1206	94.4	226	4	AAB80958 Viral pro
6	1206	94.4	226	5	Abb77789 Amino aci
7	1206	94.4	226	5	Abb05220 Hepatitis
8	1206	94.4	226	5	Aau97647 Hepatitis
9	1206	94.4	226	8	Adh77879 Hepatitis

10	1206	94.4	318	2	AAR60174	Aar60174 Chimeric
11	1206	94.4	383	1	AAP80416	Aap80416 Sequence
12	1206	94.4	406	8	ADH77843	Adh77843 Hepatitis
13	1206	94.4	406	8	ADO48159	Ado48159 Hepatitis
14	1206	94.4	406	8	ADS91504	Ads91504 Amino aci
15	1206	94.4	406	9	AEA21247	Aea21247 Nanoparti
16	1206	94.4	582	7	ADD69670	Add69670 Chimeric
17	1206	94.4	590	7	ADD69668	Add69668 Chimeric
18	1206	94.4	658	7	ADD69666	Add69666 Chimeric
19	1206	94.4	1109	7	ADD69672	Add69672 Chimeric
20	1203	94.1	226	5	ABR05221	Abb05221 Hepatitis
21	1203	94.1	226	5	Aau97648	Aau97648 Hepatitis
22	1203	94.1	236	1	AAP81992	Aap81992 Hepatitis
23	1200	93.9	226	5	AAM49180	Aam49180 HBV subty
24	1200	93.9	396	2	AAR60177	Aar60177 Chimeric
25	1199	93.8	281	1	AAP60163	Aap60163 Subtype a
26	1199	93.8	281	1	AAP60560	Aap60560 Hepatitis
27	1199	93.8	281	1	AAP60617	Aap60617 C-termina
28	1199	93.8	385	2	AAR60176	Aar60176 Chimeric
29	1198	93.7	226	9	AEA08836	Aea08836 Hepatitis
30	1197	93.7	400	2	AAR93801	Aar93801 Hepatitis
31	1197	93.7	400	6	ABR55863	AbR55863 HBV S-pro
32	1197	93.7	400	8	ADM49718	Adm49718 Hepatitis
33	1197	93.7	400	8	ADU74394	Adu74394 HBV S-pro
34	1195	93.5	281	1	AAP70294	Aap70294 Subtype a
35	1195	93.5	499	2	AAR60175	Aar60175 Chimeric
36	1193	93.3	400	4	AAG66930	Aag66930 HBV genot
37	1192	93.3	344	9	AEA08851	Aea08851 Hepatitis
38	1190	93.1	269	4	AAM48418	Aam48418 Protein #
39	1190	93.1	281	2	AAR62870	Aar62870 Hepatitis
40	1190	93.1	281	4	AAM48419	Aam48419 Protein #
41	1190	93.1	389	1	AAP60794	Aap60794 Atr-type
42	1190	93.1	843	3	AAY54044	Aay54044 Amino aci
43	1189	93.0	226	3	AAY93739	Aay93739 A wild ty
44	1188	93.0	226	3	AAM48415	Aam48415 HBV surfa
45	1187	92.9	226	9	AEA08848	Aea08848 Hepatitis

ALIGNMENTS

RESULT 1

ID AAY54045 standard; protein; 400 AA.

XX AC

XX AC AAY54045;

DT 27-MAR-2000 (first entry)

DE Amino acid sequence of a HBV large surface antigen protein.

XX HBV; HBV surface antigen-'S'-133 Oon strain (Met to Thr); DNA polymerase; large surface antigen; core protein; transactivating X protein; hepatitis vaccine; HBV infection; hepatocellular carcinoma.

OS Hepatitis B virus.

XX Key Location/Qualifiers

FT Misc-difference 14 /note= "encoded by AGC"

FT Misc-difference 174..400 /note= "these residues are specifically claimed in claim 29"

FT Misc-difference 298..320 /note= "these residues are specifically claimed in claim 23"

FT Misc-difference 298..320 /note= "these residues are specifically claimed in claim 23"

FT Misc-difference 298..320 /note= "these residues are specifically claimed in claim 23"

XX WO9966048-A1.

XX WO9966048-A1.

XX 23-DEC-1999.

XX 19-JUN-1998; 98WO-SG000046.

XX 19-JUN-1998; 98WO-SG000046.

XX 19-JUN-1998; 98WO-SG000046.

XX 19-JUN-1998; 98WO-SG000046.

XX 19-JUN-1998; 98WO-SG000046.

```
XX PA (GOVE-) GOVERNMENT REPUBLIC SINGAPORE.
XX PI Oon CJ, Lim GK, Zhao Y, Chen WN;
XX DR WPI; 2000-106104/09.
XX DR N-PSDB; AA237088.
XX XX New isolated hepatitis B virus strain, useful for, e.g. treatment of
XX PT hepatitis infection.
XX PS Claim 23; Page 39-40; 68pp; English.
XX XX
CC The present sequence is encoded by the genome of an isolated strain of
CC Hepatitis B virus designated human Hepatitis B virus (HBV) surface
CC antigen-'S'-133 Oon strain (Met to Thr). The viral genome is deposited as
CC ECCC accession numbers P97121501, P97121502 and P97121503. The nucleotide
CC sequence was isolated from hepatocellular carcinoma (HCC). The nucleotide
CC sequence encodes four overlapping proteins, which are a DNA polymerase, a
CC large surface antigen, a core protein, and a transactivating X protein.
CC The large surface antigen differs from the wild type sequence in that it
CC contains a Thr at position 133 of the wild type sequence instead of a
CC Met. The proteins are used to produce antibodies. The proteins,
CC polynucleotide and antibodies can be used for detecting the novel HBV
CC strain. The HBV polypeptides can also be used in hepatitis vaccines. The
CC HBV novel strain polypeptides can be used to identify compounds for
CC treating or preventing HBV infection or hepatocellular carcinoma
XX XX
SQ Sequence 400 AA;

Alignment Scores:
Pred. No.: 4,35e-112 Length: 400
Score: 1230.00 Matches: 224
Percent Similarity: 98.68% Conservative: 0
Best Local Similarity: 98.68% Mismatches: 2
Query Match: 96.24% Indels: 2
DB: 3 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x AAY54045 (1-400)

QY 1 ATGGAGAACACACATCAGATTCCTAGGACCCCTGCTGCTTACAGGGGGGTTTTTC 60
DB 175 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 194
QY 61 TCGTTGACAAAGATCCTCAATACCGCAGAGTCTAGATC-TGGTGGACTTCTCTCAAT 119
DB 195 SerLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 214
QY 120 TTTCTAGGGGAGCACCCACGTTCCTCGCCAAATTCGACAGTCCCACTCCCAATCA 179
DB 215 PheLeuGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 234
QY 180 CTCACCAACTCTTGTCCTCAATTTGTCCTGCTATCCTGCTGATGCTGCTGGGGTTT 239
DB 234 sSerProThrSerCysProProIleCysProGlyTyArgTrpAsnCysLeuArgArgPh 254
QY 240 TATCATATTCTCTTCATCTGCTGCTATGCTCATCTCTTGTGTTCTTCTGACTA 299
DB 254 ellellePheLeuPheIleLeuLeuLeuCysleullePheLeuLeuValLeuLeuAspTy 274
QY 300 CCAAGTATGTGCGCGTTGTCCTTACTTTCAGGAACATCAACACGACGCGGGCC 359
DB 274 rGIndlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 294
QY 360 ATGCAGACCTGCACACCTCTGCTCAAGAACTCTAGCTTCCTCTTGTGCTGTAC 419
DB 294 oCyslysthrCysThrThrProAlaGlnGlyAsnSerThrPheProSerCysCysTh 314
QY 420 AAAACCTTCGGACGGAAACTGCACTTGATTCCTCCATCCATCATCTCGGCTTTCGCAAG 479
DB 314 rlystProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaAr 334
QY 480 ATTCTATGGAGTGGGGCTCAGTCGCCGTTTCTCCGCTCAGTTTACTAGTGCCATTGT 539

Db 334 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa 354
QY 540 TCAGTGGTTCTAGGGCTTCCCCACACTGTTGGCTTTTCAGTTATATGATGATGTA 599
DB 354 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 374
QY 600 TTGGGGCGAAGTCTGTCAACATCTTGAGTCCCTTTTACCTCTATTACCAATTTTCTT 659
DB 374 rTrpGlyArgSerLeuTyAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh 394
QY 660 TTGTCTTTGGGTATACATT 678
DB 394 eCysLeuTrpValTyrlle 400

RESULT 2
AAP40310
ID AAP40310 standard; protein; 226 AA.
XX AAP40310;
XX 27-AUG-2003 (revised)
XX 25-MAR-2003 (revised)
XX 18-AUG-1992 (first entry)
XX Hepatitis virus adr-B surface antigen.
XX HBsAg; vaccine; diagnosis; HBV infection.
XX Hepatitis B virus.
XX JP59074985-A.
XX 27-APR-1984.
XX 19-OCT-1982; 82JP-00183432.
XX 19-OCT-1982; 82JP-00183432.
XX (TAKE ) TAKEDA CHEM IND LTD.
XX WPI; 1984-143231/23.
XX DNA used in prevention of infections by hepatitis virus B - comprises
XX structural gene of hepatitis virus adr-B surface antigen, coding gene and
XX at least 1 virus core antigen structural gene.
XX Disclosure; Fig 2; 13pp; Japanese.
XX The sequence is that of hepatitis virus adr-B surface antigen (HBsAg). It
XX can be used as a vaccine for the prevention of infections by hepatitis B
XX virus (HBV). See also AAP40311. (Updated on 25-MAR-2003 to correct PA
XX field.) (Updated on 27-AUG-2003 to correct OS field.)
XX SQ Sequence 226 AA;

Alignment Scores:
Pred. No.: 8,82e-110 Length: 226
Score: 1206.00 Matches: 220
Percent Similarity: 96.92% Conservative: 0
Best Local Similarity: 96.92% Mismatches: 6
Query Match: 94.37% Indels: 2
DB: 1 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x AAP40310 (1-226)

QY 1 ATGGAGAACACACATCAGATTCCTAGGACCCCTGCTGCTTACAGGGGGGTTTTTC 60
DB 1 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20
QY 61 TCGTTGACAAAGATCCTCAATACCGCAGAGTCTAGATC-TGGTGGACTTCTCTCAAT 119
DB 21 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 40
```


ID	AA898043	standard; protein; 226 AA.	Db	60	sSerProThrSerCysProProlleCysProGlyTyzArgTrpMetCysLeuArgArgPh	80
XX	AA898043;		Qy	240	TATCATATTCCTCTTCATCTGCTGCTATGCTCATCTTCTGTTGGTCTCTCGACTA	299
XX	15-AUG-2001	(first entry)	Db	80	eilePheLeuPheileLeuLeuLeuCysLeuilePheLeuLeuValLeuLeuAspIy	100
XX	Synthetic adr	type hepatitis B surface antigen SEQ ID NO:12.	Qy	300	CAAAGGTATGTTGCCGTTTCTCTACTTCCAGGAACATCAACCCAGCAGCGGGGCC	359
XX	Methanococcus;	Hepatitis B virus; Hepatitis C virus; HIV; PPIase;	Db	100	xGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr	120
KW	hepatitis B surface antigen;	hepatitis B core antigen; virus antigen;	Qy	360	ATGCAAGACCTGCACGACTCTCTGCTCAAGAAACTCTACAGTTTCCCTCTTGTGTGTGAC	419
KW	peptidyl prolyl cis-trans isomerase;	virucide; antiinflammatory;	Db	120	oCysLysThrCysThrileProAlaGlnGlyThrSerMetPheProSerCysCysTh	140
KW	anti-HIV; anti-hepatitis B;	antihepatitis-C; hepatotropic; vaccine;	Qy	420	AAACCTTCGACGGAACCTGCTGTTATTCCTCCATCCATCCCTGGCTTCGCAAG	479
XX	immunodiagnostic.		Db	140	rLysProSerAspGlyAenCysThrCysileProlleProSerTrpAlaPheAlaR	160
OS	Hepatitis B virus.		Qy	480	ATTCCTATGGAGTGGGCTCAGTCCGTTTCTCTGGCTCAGTTTACTAGTGCATTTGT	539
OS	Synthetic.		Db	160	gPheLeuTrpGlnTrpAlaSerValArgPheSerTrpLeuSerLeuValProPheVa	180
XX	JP2001066309-A.		Qy	540	TCAGTGTTCGTAGGGCTTTCCCCACATGTTGGCTTTTTCAGTTATATGATGATGTGTA	599
XX	16-MAR-2001.		Db	180	lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValilleTrpMetMetTripy	200
XX	27-SEP-1999;	99JP-00273203.	Qy	600	TTGGGGGCGAAGTCTGTACACATCTGAGTCCCTTTTACCTCTATTACCAATTTCTT	659
XX	28-DEC-1998;	98JP-00377103.	Db	200	rTrpGlyProSerLeuTyraenilleLeuSerProPheLeuProleLeuProillePhePh	220
XX	17-MAY-1999;	99JP-00136346.	Qy	660	TTGTCTTTGGGTATACATT	678
XX	(SEKI) SEKISUI CHEM IND CO LTD.		Db	220	eCysLeuTrpValTyrlle	226
XX	(KAIY-) KAIYO BIOTECHNOLOGY KENKYUSHO KK.		RESULT 5			
XX	WPI; 2001-304300/32.		AA898058			
XX	N-PSDB; AAH21782.		ID	AA898058	standard; protein; 226 AA.	
XX	Preparation of viral antigen, especially hepatitis B, hepatitis C or HIV,		XX	AC	AA898058;	
XX	for use as vaccine and in diagnostics, comprises reacting peptidyl prolyl		XX	XX	08-JUN-2001 (first entry)	
XX	cis-trans isomerase to the viral antigen.		XX	DT	Viral protein sequence #4.	
XX	Example 7; Page 21-22; 31pp; Japanese.		XX	DE	Virus antigen; molecular chaperone; immunoassay reagent.	
XX	The present invention describes a method (M) for the preparation of a		XX	OS	Unidentified.	
XX	viral antigen (I) comprising reacting (I) with peptidyl prolyl cis-trans		XX	XX	JP2001033449-A.	
XX	isomerase (PPIase). (I) has virucide, antiinflammatory, anti-HIV, anti-		XX	PD	09-FEB-2001.	
XX	hepatitis B, antihpatitis-C and hepatotropic activities. (I) is useful		XX	XX	27-SEP-1999; 99JP-00273204.	
XX	for immunodiagnosics, and as a vaccine. (I) has stronger antigenic		XX	PF	28-DEC-1998; 98JP-00377102.	
XX	activity than previously available antigens. Large quantities of viral		XX	XX	17-MAY-1999; 99JP-00136334.	
XX	antigen can be prepared by the simple process within a short culture		XX	XX	(SEKI) SEKISUI CHEM IND CO LTD.	
XX	period. The present sequence represents a synthetic adr type hepatitis B		XX	PA	(KAIY-) KAIYO BIOTECHNOLOGY KENKYUSHO KK.	
XX	surface antigen, which is given in an example from the present invention		XX	XX	WPI; 2001-230268/24.	
XX	Sequence 226 AA;		XX	DR	N-PSDB; AAP81488.	
XX	Alignment Scores:		XX	PT	Preparation of a virus antigen.	
Pred. No.:	8,82e-110	Length: 226	XX	PS	Example 7; Page 23; 33pp; Japanese.	
Score:	1206.00	Matches: 220	XX	XX	The present invention relates to a method for the preparation of a virus	
Percent Similarity:	96.92%	Conservative: 0	XX	CC	antigen. The method consists of a step of reacting a molecular chaperone	
Best Local Similarity:	96.92%	Mismatches: 6	XX	CC	with a virus antigen. The virus antigen is useful for the preparation of	
Query Match:	94.37%	Indels: 2	XX	CC	immunoassay reagents. The present sequence is a viral protein, which was	
DB:	4	Gaps: 0	XX	CC	used in the present invention	
US-10-761-006a-1_COPY_155_835 (1-681) x AA898043 (1-226)						
Qy	1	ATGGAGAACACACATCAGATTCTTAGGACCCCTGCTGTTACAGCGGGGTTTTC				60
Db	1	MetGluAenThrThrSerGlyPheLeuGlyProLeuLeuValleuGlnAlaGlyPhePhe				20
Qy	61	TGCTGACAAAGATCCTCAATACCGCAGAGTCTAGATC-TGGTGGACTTCTCTCAAT				119
Db	21	LeuLeuThrArgileLeuThrileProGlnSerLeuApsSerTrpTrpThrSerLeuAen				40
Qy	120	TTTCTAGGGGAGACCCACGTTTCTCGSCCAAAATTCGACAGTCCCAACTCCCAATCA				179
Db	41	PheLeuGlyGlyAlaProThrCys-ProGlyGlnAenSerGlnSerProThrSerAenHl				60
Qy	180	CTCACCAACCTCTTGTCTCTCAATTTGTCTGCTATCGTGTGATGTGTCTGGCGGTTT				239

XX Sequence 226 AA;
SQ
Alignment Scores:
Pred. No.: 8.82e-110 Length: 226
Score: 1206.00 Matches: 220
Percent Similarity: 96.92% Conservative: 0
Best Local Similarity: 96.92% Mismatches: 6
Query Match: 94.37% Indels: 2
DB: 4 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x AAB80958 (1-226)

QY 1 ATGAGAGACACATCAGATTCCTAGAGACCCCTGCTGTGTACAGGCGGGTTTTC 60
DB 1 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20
QY 61 TCGTTGACAGAATCCTCACAATACGACAGTCTAGACTC-TGGTGGACTTCTCTCAAT 119
DB 21 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 40
QY 120 TTTCTAGGGGAGACCCAGCTGTCTTCTGGCCAAATTCGAGTCCCAACCTCCCAATCA 179
DB 41 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 60
QY 180 CTCACCAACCTCTTGTCTCAATTTGCTCTGCTATCGCTGGAGTGTCTGGCGGTTT 239
DB 60 sSerProThrSerCysProIleCysProGlyTyArgTrpMetCysLeuArgArgPh 80
QY 240 TATCATATTCCTTCTCATCTGCTGCTATGCCTCATCTTCTGTGTCTTCTGTGACTA 299
DB 80 eilePheLeuPheIleLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTy 100
QY 300 CCAAGGTATGTGCGGCTTTGCTCTTACTTCTCAGGGAACATCAACCAAGCAGCGGGCC 359
DB 100 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 120
QY 360 ATGCAAGACTGACGACTCTGCTCAAGAAACTCTACCTTCCCTCTTGTGTCTGTCTATC 419
DB 120 oCysLysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysSth 140
QY 420 AAAACCTTCGAGCGGAACCTGCACTGTATTCCTCATCCATCCATCCTCGGCTTTCGCAAG 479
DB 140 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaAr 160
QY 480 ATTCTATGGAGTGGGCTTCAGTCCGCTTCTCTGTCTCTGTCTGTCTGTCTGTCTGTCT 539
DB 160 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa 180
QY 540 TCAGTGTTCGTAGGGCTTTCCCACTGTTTGGCTTTCAGTATATGAGATGAGTGA 599
DB 180 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 200
QY 600 TTGGGGCGGAGTCTGTACACATCTTGAGTCCCTTTTACCTTATACCAATTTCTT 659
DB 200 rTrpGlyProSerLeuTyraAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh 220
QY 660 TTGTCTTTGGGTATACATT 678
DB 220 eCysLeuTrpValIle 226

RESULT 6

ID ABB77789
ABB77789 standard; protein; 226 AA.

XX ABB77789;

XX 20-AUG-2002 (first entry)

XX Amino acid sequence of a hepatitis B virus (HBV) protein.

XX HBV; surface antigen; HBs antigen.

XX

OS Hepatitis B virus.
XX JP2002101887-A.
XX 09-APR-2002.
XX 29-SEP-2000; 2000JP-00299262.
XX 29-SEP-2000; 2000JP-00299262.
XX (SEKI) SEKISUI CHEM IND CO LTD.
XX WPI; 2002-440457/47.
XX N-PSDB; ABL59070.
XX Preparing a recombinant hepatitis B virus surface antigen and an antigen.
XX Disclosure; Page 8-9; 9pp; Japanese.

XX The specification describes a method for preparing a recombinant hepatitis B virus (HBV) surface (HBs) antigen. The recombinant HBs antigen is useful in the preparation of immunoassay reagents. The present sequence represents a HBV protein, which was used in the course of the invention

SQ Sequence 226 AA;

Alignment Scores:
Pred. No.: 8.82e-110 Length: 226
Score: 1206.00 Matches: 220
Percent Similarity: 96.92% Conservative: 0
Best Local Similarity: 96.92% Mismatches: 6
Query Match: 94.37% Indels: 2
DB: 5 Gaps: 0
US-10-761-006A-1_COPY_155_835 (1-681) x ABB77789 (1-226)

QY 1 ATGAGAGACACATCAGATTCCTAGAGACCCCTGCTGTGTACAGGCGGGTTTTC 60
DB 1 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20
QY 61 TCGTTGACAGAATCCTCACAATACGACAGTCTAGACTC-TGGTGGACTTCTCTCAAT 119
DB 21 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 40
QY 120 TTTCTAGGGGAGACCCAGCTGTCTTCTGCGCAAAATTCGAGTCCCAACCTCCCAATCA 179
DB 41 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 60
QY 180 CTCACCAACCTCTTGTCTTCTCAATTTGCTCTGCTATCGCTGATGTCTGGCGGTTT 239
DB 60 sSerProThrSerCysProIleCysProGlyTyArgTrpMetCysLeuArgArgPh 80
QY 240 TATCATATTCCTTCTCATCTGCTGCTATGCCTCATCTTCTGTGTCTTCTGTGACTA 299
DB 80 eilePheLeuPheIleLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTy 100
QY 300 CCAAGGTATGTGCGGCTTTGCTCTTACTTCTCAGGGAACATCAACCAAGCAGCGGGCC 359
DB 100 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 120
QY 360 ATGCAAGACTGACGACTCTGCTCAAGAAACTCTACCTTCCCTCTTGTGTCTGTCTATC 419
DB 120 oCysLysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysSth 140
QY 420 AAAACCTTCGAGCGGAACCTGCACTGTATTCCTCATCCATCCATCCTCGGCTTTCGCAAG 479
DB 140 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaAr 160
QY 480 ATTCTATGGAGTGGGCTTCAGTCCGCTTCTCTGTCTCTGTCTGTCTGTCTGTCTGTCT 539
DB 160 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa 180

```

QY 540 TCAGTGGTTCGTAGGGCTTCCCCACACGTGTTGGCTTTCAGTTATATGATGATGTGTA 599
Db 180 LGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 200
QY 600 TTGGGGCGGAGCTCTGTAACAACATCTTGAGTCCCTTTTACCTCTATTACCAATTTCTT 659
Db 200 rTrpGlyProSerLeuTyAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh 220
QY 660 TTGTCTTTGGGTATACATT 678
Db 220 eCysLeuTrpValTyIle 226

RESULT 7
ABB05220
ID ABB05220 standard; protein; 226 AA.
XX
AC ABB05220;
XX
DT 03-APR-2002 (first entry)
XX
DE Hepatitis B virus surface antigen protein SEQ ID NO:3.
XX
KW Hepatitis B virus; surface antigen; HBs antigen; immunoassay.
XX
OS Hepatitis B virus.
XX
PN JP2001292780-A.
XX
PD 23-OCT-2001.
XX
PF 12-APR-2000; 2000JP-00110946.
XX
PR 12-APR-2000; 2000JP-00110946.
XX
PA (SEKI ) SEKISUI CHEM IND CO LTD.
XX
WP1; 2002-126049/17.
DR N-PSDB; ABA92881.
XX
PT Preparation of a recombinant hepatitis B virus surface antigen and the
recombinant hepatitis B virus surface antigen.
XX
PS Disclosure; Page 8-9; 11pp; Japanese.
XX
CC The present invention describes a method for the preparation of a
recombinant hepatitis B virus surface antigen (HBs antigen) in which at
least 23 amino acids from the N-terminal in a HBs antigen is deleted. The
HBs antigen is useful for the preparation of reagents for immunoassay.
CC The present sequence represents a HBs antigen which is used in the
exemplification of the present invention
XX
SQ Sequence 226 AA;

Alignment Scores:
Pred. No.: 8.82e-110 Length: 226
Score: 1206.00 Matches: 220
Percent Similarity: 96.92% Conservative: 0
Best Local Similarity: 96.92% Mismatches: 6
Query Match: 94.37% Indels: 2
DB: 5 Gaps: 0

US-10-761-006a-1_copy_155_835 (1-681) x ABB05220 (1-226)

QY 1 ATGGAGAACACAACTCAGGATTCCTAGGACCCCTGCTGTGTATACAGGGGGTTTTC 60
Db 1 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnIleGlyPhePhe 20
QY 61 TCGTTGACAGAATCCTCAATACCGCAGAGTCTAGACTC-TGGTGGACTTCTCTCAAT 119
Db 21 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 40
QY 120 TTCTAGGGGGAGCACCACGTGTTCTGGCCAAATTCGACATTCGAGTCCCACTCCCAATCA 179

```

```

Db 41 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 60
QY 180 CTCACCAACCTCTTGTCTCCAAATTTGCTCGGCTATCGCTGATGTGTCTCGCGCGTTT 239
Db 60 sSerProThrSerCysProProIleCysProGlyTyArgTrpMetCysLeuArgArgPh 80
QY 240 TATCATATCTCTTCATCTCGCTGCTGCTATGCTCATCTTCTTGTGGTTCCTCTGAGACTA 299
Db 80 eIleIlePheLeuPheIleLeuLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTy 100
QY 300 CCAAGGTATGTTGCCCGTTTGTCTCTACTTCCAGGAACATCAACACACGACGCGGGCC 359
Db 100 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 120
QY 360 ATGCAAGACCTGCACGACTCTCTGCTCAAGGAAACTCTACGTTTCCCTCTTGTGTGTGTAC 419
Db 120 oCysIysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysTh 140
QY 420 AAAACCTTCGGACGGAAGAACTGACCTTGATTTCCCATCCCATCATCTCGGCTTTCGCAAG 479
Db 140 rIysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaAr 160
QY 480 ATTCTATGAGAGTGGGCTCAGTCGCTTCTCCTGGCTCAGTTTACTAGTCCCATTTGT 539
Db 160 gPheLeuTrpGlnTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValPropheVa 180
QY 540 TCAGTGGTTCGTAGGGCTTTCCTCCACATGTTTGGCTTTTTCAGTTATATGATGATGTGTA 599
Db 180 LGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 200
QY 600 TTGGGGCGGAGCTCTGTACAACATCTTGAGTCCCTTTTACCTCTATTACCAATTTCTT 659
Db 200 rTrpGlyProSerLeuTyAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh 220
QY 660 TTGTCTTTGGGTATACATT 678
Db 220 eCysLeuTrpValTyIle 226

RESULT 8
AAU97647
ID AAU97647 standard; protein; 226 AA.
XX
AC AAU97647;
XX
DT 13-AUG-2002 (first entry)
XX
DE Hepatitis B surface antigen #3.
XX
KW Hepatitis B surface antigen; immunoassay.
XX
OS Hepatitis B virus.
XX
PN JP2002112797-A.
XX
PD 16-APR-2002.
XX
PF 21-SEP-2000; 2000JP-00287232.
XX
PR 03-AUG-2000; 2000JP-00235834.
XX
PA (SEKI ) SEKISUI CHEM IND CO LTD.
XX
WP1; 2002-448759/48.
DR N-PSDB; ABK52589.
XX
PT Preparation of a recombinant hepatitis B virus surface antigen and a
hepatitis B virus surface antigen used for the preparation of immunoassay
reagents.
XX
PS Disclosure; Page 9-10; 11pp; Japanese.
XX
CC This invention relates to a novel method for preparing recombinant
hepatitis B virus surface antigen by using an antigen-determining region

```

CC containing four antigen sites of the S antigen of the hepatitis B virus
CC surface antigen and a region encoding at least four amino acids forming
CC the membrane-penetrating region. The invention also comprises a hepatitis
CC B virus surface antigen prepared using this method. The method of the
CC invention is useful for the preparation of immunoassay reagents. The
CC present sequence represents the Hepatitis B surface antigen protein #3 of
CC the invention
XX
SQ Sequence 226 AA;

Alignment Scores:
Pred. No.: 8,82e-110 Length: 226
Score: 1206.00 Matches: 220
Percent Similarity: 96.92% Conservative: 0
Best Local Similarity: 96.92% Mismatches: 6
Query Match: 94.37% Indels: 2
DB: 5 Gaps: 0

US-10-761-006a-1_COPY_155_835 (1-681) x AAU97647 (1-226)

QY 1 ATGAGAACACATCAGATTCCTAGGACCCCTGCTGTATACAGCGGGGTTTTTC 60
Db 1 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20
QY 61 TCCTTGACAGAAATCCTCACAAATACCGCAGAGTCTAGACTC-TGGTGGACTTCTCTCAAT 119
Db 21 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 40
QY 120 TTTCTAGGGGAGCAGCCAGTGTCTCGGCCAAAATTCGAGTCCCACTCCCAATCA 179
Db 41 PheLeuGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 60
QY 180 CTCACCAACCTCTTGCTCCCAATTTGCTCGGCTATCGCTGATGTCGGGGGTTT 239
Db 60 sSerProThrSerCysProProlleCysProGlyTyrArgTrpMetCysLeuArgPh 80
QY 240 TATCATATTCCTCTTCATCTGCTGTATGCTCATCTTCTGTGTGTTCTTCTGACTA 299
Db 80 eIleIlePheLeuPheIleLeuLeuCysLeuIlePheLeuValLeuLeuAspTy 100
QY 300 CCAAGGTATGTCGGGCTTCTGCTTCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 359
Db 100 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 120
QY 360 ATGCAAGACCTGCAGCAGCTCTGCTCAAGAAACTCTACGTTTCCCTCTTGTGCTGTAC 419
Db 120 oCysLysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysTh 140
QY 420 AAAACCTTCGAGCGGAACCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 479
Db 140 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAla 160
QY 480 ATTCTATGGAGTGGGCGCTCAGTCCGTTTCTCTGCTCAGTTTACTAGTGCATTGT 539
Db 160 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa 180
QY 540 TCAGTGTGCTGAGGGCTTTCCCGACTGTTTGGCTTTCAGTTATATGATGTGGTA 599
Db 180 iGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 200
QY 600 TTGGGGCGGAGTCTGTACAACTCTTGAGTCCCTTTTACCTTATTAACAATTTCTT 659
Db 200 rTrpGlyProSerLeuTyrAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh 220
QY 660 TTGCTTTTGGGTATACATT 678
Db 220 eCysLeuTrpValIle 226

RESULT 9

ADH77879

ID ADH77879 standard; protein; 226 AA.

XX

AC ADH77879;

XX 22-APR-2004 (first entry)
XX Hepatitis B virus surface antigen, HBsAgS.
XX Nanoparticle; surface antigen protein; HBsAgS; liver; drug delivery.
XX Hepatitis B virus.
XX WO2004002459-A1.
XX 08-JAN-2004.
XX 27-JUN-2003; 2003WO-JP008244.
XX 28-JUN-2002; 2002JP-00191386.
XX 27-JUN-2003; 2003JP-00183863.
XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX Kuroda S, Tanizawa K, Kondo A, Ueda M, Seno M, Tada H;
XX WPI; 2004-082876/08.
XX Hollow nanoparticle comprising protein with modified cysteine residues
XX for target specific drug delivery.
XX Example 1; Fig 2; 79pp; Japanese.
XX The present invention relates to a hollow nanoparticle, which comprises a
XX protein, e.g. hepatitis B virus surface antigen protein, with modified
XX cysteine residues. The nanoparticle specifically recognizes a target cell
XX e.g. a liver cell. The nanoparticle is useful in methods for drug
XX delivery. The present sequence is one such hepatitis B virus surface
XX antigen protein.
XX Sequence 226 AA;
Alignment Scores:
Pred. No.: 8,82e-110 Length: 226
Score: 1206.00 Matches: 220
Percent Similarity: 96.92% Conservative: 0
Best Local Similarity: 96.92% Mismatches: 6
Query Match: 94.37% Indels: 2
DB: 5 Gaps: 0
US-10-761-006a-1_COPY_155_835 (1-681) x ADH77879 (1-226)

QY 1 ATGAGAACACATCAGATTCCTAGGACCCCTGCTGTATACAGCGGGGTTTTTC 60
Db 1 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20
QY 61 TCCTTGACAGAAATCCTCACAAATACCGCAGAGTCTAGACTC-TGGTGGACTTCTCTCAAT 119
Db 21 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 40
QY 120 TTTCTAGGGGAGCAGCCAGTGTCTCGGCCAAAATTCGAGTCCCACTCCCAATCA 179
Db 41 PheLeuGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 60
QY 180 CTCACCAACCTCTTGCTCCCAATTTGCTCGGCTATCGCTGATGTCGGGGGTTT 239
Db 60 sSerProThrSerCysProProlleCysProGlyTyrArgTrpMetCysLeuArgPh 80
QY 240 TATCATATTCCTCTTCATCTGCTGTATGCTCATCTTCTGTGTGTTCTTCTGACTA 299
Db 80 eIleIlePheLeuPheIleLeuLeuCysLeuIlePheLeuValLeuLeuAspTy 100
QY 300 CCAAGGTATGTCGGGCTTCTGCTTCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 359
Db 100 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 120
QY 360 ATGCAAGACCTGCAGCAGCTCTGCTCAAGAAACTCTACGTTTCCCTCTTGTGCTGTAC 419

Db 120 oCysLysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysTh 140
QY 420 AAAACCTTCGGACGGAACCTGCACCTTGATTCCTCCATCCCATCATCTGGGCTTCGCAAG 479
Db 140 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaAr 160
QY 480 ATTCTATGAGGAGTGCGGCTCAGTCGTTCTCTGGCTCAGTTACTAGTCCCAATTGT 539
Db 160 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa 180
QY 540 TCAGTGTTCGTAGGGCTTCCTCCACCTGTTGGCTTCAGTATATGATGATCGTA 599
Db 180 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 200
QY 600 TTGGGGCGAAGTCTGTACAACTCTGAGTCCCTTTTACCTCTATTACCAATTTCTT 659
Db 200 rTrpGlyProSerLeuTyAsnIleLeuSerProPheLeuProLeuProIlePhePh 220
QY 660 TTGTCTTTGGGTATACATT 678
Db 220 eCysLeuTrpValTyIle 226

RESULT 10

AAP80416
ID AAP80174 standard; protein; 318 AA.
XX
AC AAP80174;
XX
DT 16-OCT-2003 (revised)
DT 07-APR-1995 (first entry)
XX
Chimeric protein (R3:S) used in multivalent vaccine.
XX
Chimeric; chimera; vaccine; multivalent; hepatitis B virus; HBV;
KW hepatitis; Japanese encephalitis virus; baculovirus.
XX
OS Hepatitis B virus.
OS Japanese encephalitis virus.
OS Chimeric.
XX
PN JP06205672-A.
XX
PD 26-JUL-1994.
XX
PF 19-MAR-1992; 92JP-00063699.
XX
PR 19-MAR-1992; 92JP-00063699.
XX
PA (JAPG) NIPPON ZEON KK.
PA (TOKS-) TOKYO SHINKAI KAGAKU SOGO KENKYUSHO ZH.
XX
DR WPI; 1994-275516/34.
DR N-PSDB; AAQ70155.
XX
PT Prodn. of chimeric proteins having antigenic sites from Japanese
PT encephalitis virus and hepatitis B virus surface antigens - also
PT recombinant baculovirus, useful as multivalent vaccine.
XX
PS Disclosure; Fig 1-3; 13pp; Japanese.
XX
CC The chimeric protein comprises antigenic sites from Japanese encephalitis
CC virus and Hepatitis B virus surface antigens. The protein may be used as
CC a multivalent vaccine. (Updated on 16-OCT-2003 to standardise OS field)
XX
SQ Sequence 318 AA;

Alignment Scores:
Pred. No.: 9.55e-110 Length: 318
Score: 1206.00 Matches: 220
Percent Similarity: 96.92% Conservative: 0
Best Local Similarity: 96.92% Mismatches: 6
Query Match: 94.37% Indels: 2

DB: 2 Gaps: 0
US-10-761-006A-1_COPY_155_835 (1-681) x AAP80174 (1-318)
QY 1 ATGGAGAACACAAACATCAGGATTCCTAGGACCCCTGCTCGTGTACAGCGGGGTTTTTC 60
Db 93 MetGluAenThrThrSerGlyPheLeuGlyProLeuValLeuGlnAlaGlyPhePhe 112
QY 61 TCGTTGACAGAATCCTCACAAATACCGCAGAGTCTAGATCTC-TGGTGAGCTTCTCTCAAT 119
Db 113 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAen 132
QY 120 TTTCTAGGGGAGACCCACCGTGTTCCTGGCCAAAATTTCGACAGTCCCACTCAATCA 179
Db 133 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 152
QY 180 CTCACCAACCTTGTCTCCAAATTGTCCTGGCTATCGCTGATGCTCTGCGGCGTTT 239
Db 152 sSerProThrSerCysProIleCysProGlyTyArgTrpMetCysLeuArgArgPh 172
QY 240 TATCATATTCTCTTTCATCTGCTATGCTCATCTTCTTGTGGTCTTCTCTGACTA 299
Db 172 eIleIlePheLeuPheIleLeuLeuCysLeuIlePheLeuValLeuLeuAspTy 192
QY 300 CCAAGGTATGTGCGGCTTGTCTTACTTCTCCAGGAAACATCAACACACGAGCGGGCC 359
Db 192 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 212
QY 360 ATGCAAGACCTGTCAGACTCTCTGCTCAAGGAAACTCTACGTTTCCCTCTTGTGTGTAC 419
Db 212 oCysLysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysTh 232
QY 420 AAAACCTTCGACGGAACCTGCACCTGTGATTTCCCATCCCATCATCTCGGCTTTCGCAAG 479
Db 232 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaAr 252
QY 480 ATTCTATGAGGAGTGCGGCTCAGTCCGTTTCTCCGGCTCAGTTTACTAGTCCCAATTGT 539
Db 252 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa 272
QY 540 TCAGTGTTCGTAGGGCTTCCCCACCTGTTGGCTTTCAGTTATATGATGATCGTA 599
Db 272 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 292
QY 600 TTGGGGCGAAGTCTGTACAACTCTGATCCTTGGTCCCTTTTACCTCTATTACCAATTTCTT 659
Db 292 rTrpGlyProSerLeuTyAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh 312
QY 660 TTGTCTTTGGGTATACATT 678
Db 312 eCysLeuTrpValTyIle 318

RESULT 11

AAP80416
ID AAP80416 standard; protein; 383 AA.
XX
AC AAP80416;
XX
DT 25-MAR-2003 (revised)
DT 12-SEP-1990 (first entry)
XX
DE Sequence of adr type hepatitis B virus (HBV) surface antigen (HBeAg) L
DE protein (M protein, S protein).
XX
KW Egg white lysozyme; hepatitis B virus (HBV) surface antigen (HBeAg);
KW Saccharomyces cerevisiae AH22R/PGLD LP39-Rct; L protein; M protein;
KW Saccharomyces cerevisiae L1LP39-Rct; Saccharomyces cerevisiae LP31-Rct;
KW S protein.
XX
OS Hepatitis B virus.
XX
PN EP288198-A.
XX

PD 26-OCT-1988.
 XX
 PF 13-APR-1988; 88EP-00303297.
 XX
 PR 20-APR-1987; 87JP-00098265.
 PR 12-OCT-1987; 87JP-00256885.
 PR 18-APR-1988; 88JP-00095335.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Fujisawa Y, Imai S, Miyazaki T;
 XX
 DR WPI; 1988-301233/43.
 DR N-PSDB; AAN80973.
 XX
 PT Peptide(s) having hepatitis B surface antigenicity - used as vaccine for
 PT prevention of hepatitis B virus infection and in diagnostic kits.
 XX
 PS Example; Fig 5; 31pp; English.
 XX
 CC HBsAg L protein (P39) is an env protein. When L protein genes are
 CC introduced into animal cells (CHO cells) they produce HBsAg particles
 CC containing M and S proteins. The advantage is that peptides having HBsAg
 CC antigenicity are excreted outside the cells and easily purified. Claimed
 CC is a eukaryotic cell transformed with the rDNA. Pref. the cell is a yeast
 CC cell, esp. S. cerevisiae AH22R-/pGLD LP39- RCT, LIP39-RCT or LP31-RCT.
 CC Also claimed is recombinant DNA, which is a DNA coding for a signal
 CC peptide which functions in a eukaryotic cell, and is bound to the 5'-
 CC terminal of a DNA coding for a peptide having HBsAg activity. The signal
 CC peptide may be a signal peptide of egg white lysozyme. When trypsin-like
 CC protease-producing yeast is utilized as the host, L protein and M protein
 CC may possibly be degraded by the protease; therefore it is desirable that
 CC the genes are altered so that the 48th arginine residue from the N-
 CC terminal of M protein or a peptide containing the residue (preferably the
 CC 44-49 peptide) may be deleted. Preferable genes include L protein gene
 CC (altered) coding for the amino acid sequence 1-383 shown in AAN80973.
 CC (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to
 CC correct PA field.)
 XX
 SQ Sequence 383 AA;

Alignment Scores:
 Pred. No.: 9.98e-110 Length: 383
 Score: 1206.00 Matches: 220
 Percent Similarity: 96.92% Conservative: 0
 Best Local Similarity: 96.92% Mismatches: 6
 Query Match: 94.37% Indels: 2
 DB: 1 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x AAP80416 (1-383)

QY 1 ATGGAGAACACAAATCAGGATTCCTAGGACCCCTGCTCGTGTACAGGCGGGTTC 60
 DB 158 MetGluAsnThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 177
 QY 61 TCGTTGACAGAAATCCCAATACCCAGAGTCTAGACTC-TGGTGGACTTCTCTCAAT 119
 DB 178 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 197
 QY 120 TTTCTAGGGGAGACCCAGTGTCTGGCCAAATTCGACGTCGCCACCTCCATCA 179
 DB 198 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnH 217
 QY 180 CTACCAACCTCTGTCTCCAAATTCCTGGCTATCGTGGATGTCTGGCGGCTTT 239
 DB 217 sSerProThrSerCysProPheLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 237
 QY 240 TATCATATTCCTCTTCATCTCGTGTATGCTCATCTTCTGTGTGTTCTTCTGGACTA 299
 DB 237 eillePheLeuPheilleLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 257
 QY 300 CCAGGTATGTCCT 359

DB 257 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 277
 QY 360 ATGCAAGACCTGCACGACTCTGCTCAAGAACTCTAGCTTCCTCTCTGTTGCTGTAC 419
 DB 277 oCysIysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysCys 297
 QY 420 AAAAACTTCGGACGGAACCTGCACCTGTATTCCCATCCCATCATCTCGGCTTTCGCAAG 479
 DB 297 rIysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAla 317
 QY 480 ATTCTATATGGAGTGGGCTCAGTCCGTTTCTCTCGCTCAGTTTACTAGTGCATTTGT 539
 DB 317 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValProPheVa 337
 QY 540 TCAAGTGTTCGTAAGGCTTTCCTCCCACTGTTGGCTTTCAGTTATATGATGATGGTA 599
 DB 337 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrp 357
 QY 600 TTGGGGCGGAGTCTGTACAACTTTGAGTCCCTTTTACCTCTATTACCAATTTTCTT 659
 DB 357 rTrpGlyProSerLeuTrpAsnIleLeuSerProPheLeuProLeuLeuProIlePhe 377
 QY 660 TTGCTTTTGGGTATACATT 678
 DB 377 eCysLeuTrpValTyIle 383
 RESULT 12
 ADH77843
 ID ADH77843 standard; protein; 406 AA.
 AC ADH77843;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE Hepatitis B virus surface antigen, HBsAgL, SEQ ID 2.
 XX
 KW Nanoparticle; surface antigen protein; HBsAgL; liver; drug delivery.
 XX
 OS Hepatitis B virus.
 XX
 PN WO2004002459-A1.
 XX
 PD 08-JAN-2004.
 XX
 PF 27-JUN-2003; 2003WO-JP008244.
 XX
 PR 28-JUN-2002; 2002JP-00191386.
 PR 27-JUN-2003; 2003JP-00183863.
 XX
 PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
 XX
 PI Kuroda S, Tanizawa K, Kondo A, Ueda M, Seno M, Tada H;
 XX
 DR WPI; 2004-082876/08.
 DR N-PSDB; ADH77842.
 XX
 PT Hollow nanoparticle comprising protein with modified cysteine residues
 PT for target specific drug delivery.
 XX
 PS Example 1; SEQ ID NO 2; 79pp; Japanese.
 XX
 CC The present invention relates to a hollow nanoparticle, which comprises a
 CC protein, e.g. hepatitis B virus surface antigen protein, with modified
 CC cysteine residues. The nanoparticle specifically recognizes a target cell
 CC e.g. a liver cell. The nanoparticle is useful in methods for drug
 CC delivery. The present sequence is one such hepatitis B virus surface
 CC antigen protein.
 XX
 SQ Sequence 406 AA;
 Alignment Scores:
 Pred. No.: 1.01e-109 Length: 406
 Score: 1206.00 Matches: 220

Percent Similarity:	96.92%	Conservative:	0
Best Local Similarity:	96.92%	Mismatches:	6
Query Match:	94.37%	Indels:	2
DB:	8	Gaps:	0
US-10-761-006A-1_COPY_155_835 (1-681) x ADH77843 (1-406)			
QY	1	ATGGAGAACACATCAGGATTCCTAGGACCCCTGCTGCTGTATACAGCGGGGTTTTTC	60
Db	181	MetGluAenThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe	200
QY	61	TCGTTGACAGAAATCCTCACAATACCGAGAGCTAGACTC-TGGTGGACTTCTCTCAAT	119
Db	201	LeuLeuThrArgIleLeuThrIleProGlnSerLeuAsePserTrpTrpThrSerLeuAse	220
QY	120	TTTCTAGGGGAGACCCACGTTGCTCGCCAAAATTCGCAAGTCCCAACCTCCAATCA	179
Db	221	PheLeuGlyGlyAlaProThrCys-ProGlyGlnAseSerGlnSerProThrSerAseHi	240
QY	180	CTCACCACCTCTTGTCTCCAAATTTGCTCGCTGATGTGTCTCGCGGTTT	239
Db	240	sSerProThrSerCysProPheCysProGlyTyArgTrpMetCysLeuAseArgPh	260
QY	240	TATCATATTCCTTTCATCTGCTGCTATGCTCTCATCTTCTTGTGGTCTTCTGACTA	299
Db	260	eileilePheLeuPheIleLeuLeuLeuCysLeuIlePheLeuLeuValLeuLeuAsePty	280
QY	300	CAAAGTATGTGGCGTTGCTCTACTTCCAGGAACATCAACACCGACGCGGGCC	359
Db	280	rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr	300
QY	360	ATGCAAGACTCGACGACTCTCTGCTCAAGAAACTCTAGTTTCCCTCTTGTGTCTGTAC	419
Db	300	oCysIysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysTh	320
QY	420	AAACCTTCGGAGGAACTGCACTGTATTCCTATCCATCCATCATCTCGGCTTCGCAAG	479
Db	320	rIysProSerAsePglyAseCysThrCysIleProIleProSerSerTrpAlaPheAlaR	340
QY	480	ATTCTATGGAGTGGGCTCAGTCCGTTTCTCTGGCTCAGTTACTAGTCCATTGT	539
Db	340	gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa	360
QY	540	TCAGTGGTTCTGAGGCTTTCCCCACTGTTTGGCTTTTCAGTTATATGATGATGTGTA	599
Db	360	lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy	380
QY	600	TTGGGGGCGAAGTCTGTACAACTCTGAGTCCCTTTTACCTCTATTAACCAATTTCTT	659
Db	380	rTrpGlyProSerLeuTyAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh	400
QY	660	TTGTCTTTGGGTATACATT	678
Db	400	eCysLeuTrpValTyrlle	406
RESULT 13			
ADO48159			
ID	ADO48159	standard; protein; 406 AA.	
XX	AC	ADO48159;	
XX	DT	26-AUG-2004 (first entry)	
XX	DE	Hepatitis B virus pre-S1 protein gene SeqID2.	
XX	KW	hollow nanoparticle; particle forming; capsid structure;	
XX	KW	hepatitis B virus; surface antigen protein; hepatocyte;	
XX	KW	biological recognition molecule; beta cell phosphorus;	
XX	KW	basic fibroblast growth factor; bFGF; inner core antigen protein; pre-S1;	
XX	OS	Hepatitis B virus.	
XX	DB		

PN	WO2004047812-A1.		
XX	10-JUN-2004.		
XX	25-NOV-2003; 2003WO-JP015003.		
XX	22-NOV-2002; 2002JP-00339925.		
XX	(NISC-) JAPAN SCI & TECHNOLOGY AGENCY.		
XX	Kuroda S, Tanizawa K, Kondo A, Ueda M, Seno M;		
XX	WPI; 2004-450201/42.		
XX	N-PSDB; ADO48158.		
XX	Hollow nanoparticle, for transferring substance specifically into target		
XX	cell, comprises first protein capable of recognizing specific cells e.g.		
XX	hepatocyte, having particle forming ability, and second protein forming		
XX	capsid structure.		
XX	Disclosure; SEQ ID NO 2; 68pp; Japanese.		
XX	This invention relates to a novel hollow nanoparticle comprising a first		
XX	protein capable of recognising specific cells and having particle forming		
XX	ability, and a second protein which forms a capsid structure on the first		
XX	protein. The first protein is preferably the hepatitis B virus surface		
XX	antigen protein, which is capable of modifying the hepatocyte recognising		
XX	site in the hepatitis B virus surface antigen protein of another		
XX	biological recognition molecule, to beta cell phosphorus or a basic		
XX	fibroblast growth factor (bFGF). The second protein is a hepatitis B		
XX	virus inner core antigen protein. The invention is useful for treating a		
XX	disease and for transferring a substance or drug specifically into a		
XX	target cell or tissue. The invention enables specific and efficient		
XX	transfer of a substance into a target cell or tissue, which can be		
XX	produced in a stable yield. The present sequence is that of the Hepatitis		
XX	B virus pre-S1 protein gene sequence which is related to the invention.		
XX	Sequence 406 AA;		
Alignment Scores:			
Pred. No.:	1,01e-109	Length:	406
Score:	1206.00	Matches:	220
Percent Similarity:	96.92%	Conservative:	0
Best Local Similarity:	96.92%	Mismatches:	6
Query Match:	94.37%	Indels:	2
DB:	8	Gaps:	0
US-10-761-006A-1_COPY_155_835 (1-681) x ADO48159 (1-406)			
QY	1	ATGGAGAACACATCAGGATTCCTAGGACCCCTGCTGCTGTATACAGCGGGGTTTTTC	60
Db	181	MetGluAenThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe	200
QY	61	TCGTTGACAGAAATCCTCACAATACCGAGAGCTAGACTC-TGGTGGACTTCTCTCAAT	119
Db	201	LeuLeuThrArgIleLeuThrIleProGlnSerLeuAsePserTrpTrpThrSerLeuAse	220
QY	120	TTTCTAGGGGAGACCCACGTTGCTCGCCAAAATTCGCAAGTCCCAACCTCCAATCA	179
Db	221	PheLeuGlyGlyAlaProThrCys-ProGlyGlnAseSerGlnSerProThrSerAseHi	240
QY	180	CTCACCACCTCTTGTCTCCAAATTTGCTCGCTGATGTGTCTCGCGGTTT	239
Db	240	sSerProThrSerCysProPheCysProGlyTyArgTrpMetCysLeuAseArgPh	260
QY	240	TATCATATTCCTTTCATCTGCTGCTATGCTCTCATCTTCTTGTGGTCTTCTGACTA	299
Db	260	eileilePheLeuPheIleLeuLeuLeuCysLeuIlePheLeuLeuValLeuLeuAsePty	280
QY	300	CAAAGTATGTGGCGTTGCTCTACTTCCAGGAACATCAACACCGACGCGGGCC	359
Db	280	rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr	300

QY 360 ATGCAAGACCTGACGACCTCTGCTCAAGGAACCTCTAGCTTCCCTCTGTTGCTGTAC 419
 Db 300 oCysLysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysTh 320
 QY 420 AAAACCTTCGGACGGAACTGCACTTGATTCCTCCATCCCATCTCTGGGCTTCGCAAG 479
 Db 320 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAla 340
 QY 480 ATTCCTATGAGAGTGGGCTCAGTCCGTTTCTCTGGCTCAGTTACTAGTGCATTGT 539
 Db 340 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValPropheVa 360
 QY 540 TCAGTGGTTCGTAGGCTTTCCCACTGTTTGGCTTTCAGTTATATGATGATGGTA 599
 Db 360 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrp 380
 QY 600 TTGGGGCGGAAGTCTGTACAACTCTTGAGTCCCTTTTACCTCTATTACCAATTTCTT 659
 Db 380 rTrpGlyProSerLeuTrpAsnIleLeuSerProPheLeuProLeuProIlePhePh 400
 QY 660 TTGCTTTTGGGTATACATT 678
 Db 400 eCysLeuTrpValTyIle 406

RESULT 14

ADS91504

ID ADS91504 standard; protein; 406 AA.

AC ADS91504;

DT 16-DEC-2004 (first entry)

DE Amino acid sequence of a HBV surface antigen protein.

KW haemophilia; nanoparticle; HBV; surface antigen;

KM blood coagulation factor VIII; haemostatic; gene therapy.

OS Hepatitis B virus.

XX WO2004082720-A1.

XX 30-SEP-2004.

XX 17-MAR-2004; 2004WO-JP003560.

XX 17-MAR-2003; 2003JP-00071788.

XX (BEAC-) BEACLE INC.

XX (VIBV-) VIB VLAMS INTERUNIVERSITAIR INST BIOTEC.

XX (COLL-) COLLEN RES FOUND VZW ONDERWIJZEN NAVORSI.

XX Ueda M, Kuroda S, Tanizawa K, Senoo M, Kondo A;

XX Vandendriessche T, Chuah M;

XX WPI; 2004-699721/68.

XX N-PSDB; ADS91503.

XX Agent useful for treating hemophilia, comprises gene encapsulated in

XX hollow nanoparticle, which is obtained by expressing protein that is

XX capable of forming particles in eukaryotic cell.

XX Example; SEQ ID NO 2; 33pp; Japanese.

XX The specification describes a haemophilia therapeutic agent, which
 CC comprises a gene encapsulated in a hollow nanoparticle, which is obtained
 CC by expressing a protein that is capable of forming particles in a
 CC eukaryotic cell. The protein that forms the particle is a hepatitis B
 CC virus (HBV) surface antigen protein. The gene is a blood coagulation
 CC factor VIII (IX) gene. Therapeutic agents of the invention are useful for
 CC treating haemophilia. They are efficiently transferred into liver cells
 CC with little risk of side effects. The present sequence represents a HBV
 CC surface antigen protein, which may be used to produce therapeutic agents
 CC of the invention.

XX Sequence 406 AA;
 SQ Alignment Scores:
 Pred. No.: 1,01e-109 Length: 406
 Score: 1206.00 Matches: 220
 Percent Similarity: 96.92% Conservative: 0
 Best Local Similarity: 96.92% Mismatches: 6
 Query Match: 94.32% Indels: 2
 DB: 8 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x ADS91504 (1-406)

QY 1 ATGCGAACAACATCAGGATTCCTAGGACCCCTGCTGCTTACAGCGGGGTTTTTC 60
 Db 181 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValGlnAlaGlyPhePhe 200
 QY 61 TCGTTGACAAAGATCCTCAATACCGAGAGTCTAGATC-TGGTGGACTTCTCTCAAT 119
 Db 201 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 220
 QY 120 TTTTAGGGGAGACCCACGTTCTCGGCCAAATTCGCAAGTCCCAACCTCCCAATCA 179
 Db 221 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 240
 QY 180 CTCACCAACCTCTGTCCTCCAAATTTGCTCCTGCTATCGCTGATGTGTCTGGCGGTTT 239
 Db 240 sSerProThrSerCysProIleCysProGlyTyArgTrpMetCysLeuArgArgPh 260
 QY 240 TATCATATTCTTTCATCTGCTGCTATGCTCATCTTCTTGTGTTCTTCTGACTA 299
 Db 260 eillelPheLeuPheilleLeuLeuLeuCysleullePheLeuLeuValLeuLeuAspTy 280
 QY 300 CCAAGGTATGTGCCCGTTTGTCTTACTTCCAGGAACATCAACCAACGACGCGGGCC 359
 Db 280 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrSerThrGlyPr 300
 QY 360 ATGCAAGACCTGACGACTCTGCTCAGCAAACTCTAGCTTCCCTCTGTTGCTGTAC 419
 Db 300 oCysLysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysTh 320
 QY 420 AAAACCTTCGGACGGAACTGCACTTGATTCCTCCATCCCATCTCTGGGCTTCGCAAG 479
 Db 320 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAla 340
 QY 480 ATTCCTATGAGAGTGGGCTCAGTCCGTTTCTCTGGCTCAGTTACTAGTGCATTGT 539
 Db 340 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValPropheVa 360
 QY 540 TCAGTGGTTCGTAGGCTTTCCCACTGTTTGGCTTTCAGTTATATGATGATGGTA 599
 Db 360 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrp 380
 QY 600 TTGGGGCGGAAGTCTGTACAACTCTTGAGTCCCTTTTACCTCTATTACCAATTTCTT 659
 Db 380 rTrpGlyProSerLeuTrpAsnIleLeuSerProPheLeuProLeuProIlePhePh 400
 QY 660 TTGCTTTTGGGTATACATT 678
 Db 400 eCysLeuTrpValTyIle 406

RESULT 15

AEA21247

ID AEA21247 standard; protein; 406 AA.

XX AC AEA21247;

XX AC AEA21247;

DT 11-AUG-2005 (first entry)

DE Nanoparticle-associated protein #1.

XX nanoparticle; Gene therapy; drug delivery; pharmaceutical.

OS Unidentified.
XX WO2005049824-A1.
XX 02-JUN-2005.
XX 19-NOV-2004; 2004WO-JP017282.
XX 21-NOV-2003; 2003JP-00392649.
XX (NISC-) JAPAN SCI & TECHNOLOGY AGENCY.
XX Seno M, Tada H, Kuroda S, Tanizawa K, Kondo A, Ueda M;
XX WPI; 2005-396109/40.
XX N-PSDB; AEA21246.
XX Hollow nanoparticle useful in preparation of medical agent for treating
PT disease, made of protein capable of forming particles, and presents
PT biological structure-recognition sites on both N- and C-terminals of
PT protein.
XX Disclosure; Page; 54pp; Japanese.
XX The invention relates to a hollow nanoparticle (I) made of protein
CC capable of forming particles and presents biological structure-
CC recognition sites on its surface, where the biological structure-
CC recognition sites are provided at both the N- and C-terminals of the
CC protein and recognize different chemical structure of biological origin
CC from each other. Also described are the following: production of (I); and
CC a medical agent (II) comprising a cell introduced with a substance and
CC sealed by (I). In (I), the biological structure recognition site is a
CC target structure recognition site which recognizes a cell that serves as
CC a target. The biological structure recognition site functions as an
CC identification region for recognizing (I) specifically. The biological
CC structure recognition site is a receptor of host cell recognition
CC structure derived from virus e.g. hepatocyte recognition site of
CC hepatitis B virus surface antigenic protein, antigenic determinant e.g.
CC tag sequence such as Strep-tagII, His-tag, HA-tag or FLAG-tag, where one
CC or more tag sequences are combined, or ligand e.g. cell growth factor or
CC ZZ-tag. The cell growth factor is epidermal growth factor fibroblast
CC growth factor. The biological structure recognition site combines a tag
CC sequence and ligand. The protein which has particle formation ability is
CC a surface antigen protein derived from virus, preferably hepatitis B
CC virus surface antigen protein. The antigenic determinant and/or ligand
CC are provided as the identification region for recognizing (I),
CC specifically. The biological structure recognition site is introduced
CC into the protein by substituting or adding amino acid sequence of the
CC protein capable of forming particle at any one of the terminal. In (II),
CC the substance introduced into the cell is a compound comprising a gene
CC having pharmacological effect. (I) is useful in the preparation of a
CC medical agent which is useful for treating a disease. (I) is useful as a
CC reagent e.g. nanocapsule in pharmaceutical industry for drug delivery.
CC (I) can be purified efficiently and can be introduced into a cell for
CC preparing a drug. The present sequence represents a nanoparticle-
CC associated protein of the invention. Note: This sequence is not shown in
CC the specification but was obtained in electronic format directly from
CC WIPO at ftp.wipo.int/pub/published_pct_sequences/02.06.2005/.
XX SQ Sequence 406 AA;

Alignment Scores:
Pred. No.: 1,01e-109 Length: 406
Score: 1206.00 Matches: 220
Percent Similarity: 96.92% Conservative: 0
Best Local Similarity: 96.92% Mismatches: 6
Query Match: 94.37% Indels: 2
DB: 9 Gaps: 0
US-10-761-006A-1_COPY_155_835 (1-681) x AEA21247 (1-406)
QY 1 ATGGAGAACACAAATCAGAGATTCCTAGACCCCTCTCGTGTATACGGCGGGTTTTC 60
|||||

Search completed: December 27, 2005, 20:51:57
Job time : 187.238 secs

Db 181 MetGluAsnThrThrSerGlyPheLeuGlyProLeuValLeuGlnAlaGlyPhePhe 200
QY 61 TCGTTGACAGAAATCTCACAATACCGCAGAGTCTAGATC-TGGTGGACTTCTCTCAAT 119
|||||
Db 201 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 220
|||||
QY 120 TTTCTAGGGGAGCAGCACCACGTTCTCGGCCAAAAATTCGAGTCCCCCAACTCCAATCA 179
|||||
Db 221 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 240
|||||
QY 180 CTCACCAACCTCTTGTCTCTCAATTTGTCTGGCTATCGCTGGATGTCTTCGCGCGTTT 239
|||||
Db 240 sSerProThrSerCysProProIleCysProGlyTyrArgTrpMetCysLeuArgArgPh 260
|||||
QY 240 TATCATATTCCTCTTCATCTGCTGCTATGCTCATCTTCTTGTGGTCTTCTCGACTA 299
|||||
Db 260 ellelePheLeuPheleleLeuLeuLeuCysleullePheLeuLeuValLeuLeuAspTy 280
|||||
QY 300 CCAAGGTATGTTGCCCGTTTGTCTCTACTTCCAGGAAACATCAACACCAGCAGCGGGCC 359
|||||
Db 280 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 300
|||||
QY 360 ATGCAAGACTGTCAGCACTCTCTGCTCAAGAAACTCTAGCTTCCCTCTTGTGTCTGTAC 419
|||||
Db 300 oCysLysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysCysTh 320
|||||
QY 420 AAAACCTTCGGACGGGAAACTGCACCTTGATTTCCCATCCATCATCTCGGCTTTCGCAAG 479
|||||
Db 330 rLysProSerAspGlyAsnCysThrCysIleProleProSerSerTrpAlaPheAlaAr 340
|||||
QY 480 ATTCTATGGGAGTGGGCTCAGTCCGTTTCTCTGGCTCAGTTTACTAGTGCCATTGT 539
|||||
Db 340 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValProPheVa 360
|||||
QY 540 TCAGTGGTTCGTAGGGCTTCCCCCAGCTGTTGGCTTTCAGTTATATGATGATGTGTA 599
|||||
Db 360 IGlntPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 380
|||||
QY 600 TTGGGGGGGAGTCTGTACACATCTTCAGTCCCTTTTACCTCTATTACCAATTTTCTT 659
|||||
Db 380 rTrpGlyProSerLeuTyAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh 400
|||||
QY 660 TTGCTTTTGGGTATACATT 678
|||||
Db 400 eCysLeuTrpValTyIle 406

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 1, 2005, 23:16:00 ; Search time 94.5 Seconds
(without alignments)
5625.446 Million cell updates/sec

Title: US-10-761-006A-1
Perfect score: 5907
Sequence: 1 CTCACACATTCACCAAG.....CTCAGGCCACGACGTGGA 3215

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-O=/cgn2_1/USPTO.spool.p/US10761006/runat_01122005_113953_26895/app.query.fasta_1.3399
-DB=Issued Patents AA -QFMT=fastan -SURFIX=n2p.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODES=LOCAL -OUTFMT=ptc -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10761006 @CGN 1 1.140 @runat_01122005_113953_26895 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DSV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5 COMB.pcp.*
2: /cgn2_6/ptodata/1/iaa/6 COMB.pcp.*
3: /cgn2_6/ptodata/1/iaa/7 COMB.pcp.*
4: /cgn2_6/ptodata/1/iaa/8 COMB.pcp.*
5: /cgn2_6/ptodata/1/iaa/9 COMB.pcp.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2927	49.6	843	2	US-09-719-528A-2
2	2927	49.6	843	2	US-10-209-264-2
3	2807	47.5	843	2	US-08-591-502B-45
4	2807	47.5	843	2	US-08-591-502B-59
5	2788	47.2	843	2	US-08-591-502B-47
6	2773	46.9	843	2	US-08-591-502B-46
7	2759	46.7	843	2	US-08-591-502B-48
8	2759	46.7	843	1	US-08-416-950-11
9	2759	46.7	845	1	US-08-469-830-11
10	2741	46.4	845	1	US-08-591-502B-11
11	2734.5	46.3	842	2	US-08-591-502B-51
12	2729.5	46.2	842	2	US-08-591-502B-50

13	2652	44.9	832	2	US-08-591-502B-61
14	2646	44.8	832	2	US-08-591-502B-62
15	2645	44.8	843	2	US-08-591-502B-56
16	2643	44.7	832	2	US-08-591-502B-60
17	2643	44.7	832	2	US-08-591-502B-55
18	2637	44.6	832	2	US-08-591-502B-64
19	2633	44.6	843	2	US-08-591-502B-53
20	2615	44.3	843	2	US-08-591-502B-54
21	2608	44.2	845	2	US-08-591-502B-57
22	2607	44.1	832	2	US-08-591-502B-63
23	2607	44.1	845	2	US-08-591-502B-58
24	2607	44.1	845	6	5196194-17
25	2179	36.9	730	2	US-08-591-502B-49
26	1496	25.3	400	2	US-09-719-528A-3
27	1496	25.3	400	2	US-10-209-264-3
28	1427	24.2	277	6	5164485-2
29	1423	24.1	281	2	US-09-724-969-10
30	1423	24.1	281	2	US-09-724-969-10
31	1423	24.1	281	2	US-09-724-852-10
32	1365.5	23.1	481	2	US-08-591-502B-52
33	1359	23.0	281	1	US-08-105-483-214
34	1359	23.0	281	1	US-08-709-209-214
35	1359	23.0	281	1	US-08-458-101-214
36	1359	23.0	389	1	US-08-105-483-216
37	1359	23.0	389	1	US-08-105-483-219
38	1359	23.0	389	1	US-08-709-209-216
39	1359	23.0	389	1	US-08-709-209-219
40	1359	23.0	389	1	US-08-458-101-216
41	1359	23.0	389	1	US-08-458-101-219
42	1354	22.9	281	2	US-09-247-890-12
43	1354	22.9	281	2	US-09-724-969-12
44	1354	22.9	281	2	US-09-724-852-12
45	1354	22.9	281	2	US-09-721-480-3

ALIGNMENTS

RESULT 1
US-09-719-528A-2
; Sequenced, Application US/09719528A
; Patent No. 6558675
; GENERAL INFORMATION:
; APPLICANT: Oon, Chong Jin
; Lim, Gek Keow
; Zhao, Yi
; Chen, Wei Ning

TITLE OF INVENTION: A MUTANT HUMAN HEPATITIS B VIRAL STRAIN AND
USSES THEREOF
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ladas & Parry
STREET: 26 West 61 Street
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10023

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA: US/09/719,528A
APPLICATION NUMBER: US/09/719,528A
FILING DATE: 30-Apr-2001
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/SG98/00046
FILING DATE: 19-JAN-1998
ATTORNEY/AGENT INFORMATION:
NAME: Mass, Clifford J.
REGISTRATION NUMBER: 30,086
REFERENCE/DOCKET NUMBER: U-013109-7
TELECOMMUNICATION INFORMATION:

```
;
; TELEPHONE: (212) 708-1800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 843 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-719-528A-2

Alignment Scores:
Pred. No.: 5,09e-259 Length: 843
Score: 2927.00 Matches: 540
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 49.55% Indels: 0
DB: 2 Gaps: 0

US-10-761-006A-1 (1-3215) x US-09-719-528A-2 (1-843)

QY 1 CTCCACAACTTCCACCAAGCTCTGCTAGATCCCAAGGGTGAGGGCCCTATATTTTCCTGC 60
DB
DB 304 LeuHisAsnIleProProSerSerAlaArgSerGlnGlyGluGlyProIlePheSerCys 323
QY 61 TGGTGGCTCCAGTTCCGGAAAGTAGTAAACCTGTTCCGACTACTGCTCTCCCATATCGTC 120
DB
DB 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerHisIleVal 343
QY 121 AATCTTCTCGAGGACTGGGGCCCTCGACCCGACCAACATGAGACACCAACATCAGATTCT 180
DB
DB 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
QY 181 AGGACCCCTGCTCGTGTGTACAGGGGGGCTTTTCTCGTTGACAAGATCTCACAATACC 240
DB
DB 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
QY 241 GCAGAGCTTAGACTGTGTGACTTCTCTCAATTTTCTAGGGGAGCACCACCGTGTCC 300
DB
DB 384 AlaGluSerArgLeuTrpTrpThrSerLeuAsnPheLeuGlyAlaProThrCysSer 403
QY 301 TGGCCMAATTCGGAGTCCCACTCCAACTCAACACTCACCACCTCTGTCTCCCAATTG 360
DB
DB 404 TrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeu 423
QY 361 TCCTGGCTATCGCTGATGTCTCGCGGCTTTTATCATATTTCTCTTCATCTCTGCTGCT 420
DB
DB 424 SerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAla 443
QY 421 ATGCTCTCATCTTCTGTTGGTCTTCTGGAATAACAAGTATGTTGGCCGTTTGTCTCT 480
DB
DB 444 MetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSer 463
QY 481 ACTTCCAGACATCAACACACGACGGGCGCATGCAAGACCTGCACGACTCTCTGCTCA 540
DB
DB 464 ThrSerArgAsnIleAsnHisGlnHisGlyAlaMetGlnAspLeuHisAspSerCysSer 483
QY 541 AGGAACTCTACGTTTCCCTCTTGTGTCTGTACAAACCTTCGACGGAAACTGCACCTG 600
DB
DB 484 ArgLysLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLeu 503
QY 601 TATTCCTCATCCATCATCTCGGCTTCGCAAGATTCCTATGGAGTGGGCGCTCAGTCCG 660
DB
DB 504 TyrSerHisProIleLeuLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPro 523
QY 661 TTTCTCTGGCTCAGTTTACTAGTGGCATTTTGTTCAGTGGTTCGTAGGGCTTTCCCCAC 720
DB
DB 524 PheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHis 543
QY 721 TGTGTGGCTTTTCAGTTATATGATGATGTGTATTTGGGGCGAAGTCTGTACAACATCTT 780
DB
DB 544 CysLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLeu 563
QY 781 GAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTTGGGTATACATTTAAACCCCT 840
```

```
DB 564 GluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPro 583
QY 841 AATAAAACCAACGTTGGGCTACTCCCTTAACTTCATCGGATATGTAATGGAGTTGG 900
DB
DB 584 AsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTrp 603
QY 901 GGTACTTTTACCGCAGGAACATATTGTACTAAACCTCAAGCAATGTTTTTCGAAAACTGCCT 960
DB
DB 604 GlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPro 623
QY 961 GTAAATAGACTATTGATTTGAAAGATGTCAAAGAAATGTGGGTCTTTTGGGCTTTGCT 1020
DB
DB 624 ValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAla 643
QY 1021 GCCCTTTTACACAAATGTGGCTATCCCTTGCCTTGTATGCTTTATATGATGATATACATCT 1080
DB
DB 644 AlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSer 663
QY 1081 AAGCAGGCTTTCACTTTCTCGCCAACTTACAAGGCCCTTTCTGTATAACAATATCTCAAC 1140
DB
DB 664 LysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAsn 683
QY 1141 CTTTACCCCGTTGCCCGCAACGGTCCGGTCTCTGCAAGTGTGTGTGACGCAACCCCC 1200
DB
DB 684 LeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPro 703
QY 1201 ACTGATGGGGCTGGCCATAGCCCATCAGCGCATGCTGGACCTTTCTGGCTCCCTCTG 1260
DB
DB 704 ThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetAlaGlyThrPheLeuAlaProLeu 723
QY 1261 CCGATCCATACTCGGGAACCTCTAGCAGCTTGTCTCGCAGCCGCTCTGGAGCAAAA 1320
DB
DB 724 ProfileHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys 743
QY 1321 CTTATCGGAACCGCAAACTCTGTTGTCTCTCTCGGAAATACACCTCTCTTTCAT 1380
DB
DB 744 LeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLeu 763
QY 1381 CTAGGGTGTCTGCCAATCTGGATCTCGGGGCGCTCTGGGAGCTCTTGTCTACGTCCTCGCG 1440
DB
DB 764 LeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAla 783
QY 1441 CTGAATCCCGCGGACGACCCGCTCTCGGGGCGCTTGGGGCTTACCGCTCCCTCTTTCAT 1500
DB
DB 784 LeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHis 803
QY 1501 CTGCGGTTCCGGCGACACCGGGCGCACTCTCTTTTACCGGGTCTCCCGTATGTGCCT 1560
DB
DB 804 LeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProTyrValPro 823
QY 1561 TCTCATCTCGCGGACCGGTGCACTTCGCTTCCCTCTGACCTCTGACGCTCGCATGAGACCCG 1620
DB
DB 824 SerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaIleTrpArgPro 843

RESULT 2
US-10-209-264-2
; Sequence 2, Application US/10209264
; Patent No. 6781142
; GENERAL INFORMATION:
; APPLICANT: Oon, Chong Jin
; Lim, Gek Keow
; Zhao, Yi
; Chen, Wei Ning
; TITLE OF INVENTION: A MUTANT HUMAN HEPATITIS B VIRAL STRAIN AND
; USES THEREOF
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESS: Ladas & Parry
; STREET: 26 West 61 Street
; CITY: New York
; STATE: New York
; COUNTRY: USA
```

```

;
; ZIP: 10023
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; FILING DATE: 31-Jul-2002
; FILING DATE: 19-Jan-1998
; CLASSIFICATION DATA:
; APPLICATION NUMBER: PCT/SG98/00046
; PRIORITY NUMBER: 30,086
; ATTORNEY/AGENT INFORMATION:
; NAME: Mass, Clifford J.
; REGISTRATION NUMBER: 013109-7
; REFERENCE/DOCKET NUMBER: U-013109-7
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 708-1800
;
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 843 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```

```

US-10-209-264-2
Alignment Scores:
Pred. No.: 5,09e-259 Length: 843
Score: 2927.00 Matches: 540
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 49.55% Indels: 0
DB: 2 Gaps: 0

```

```

US-10-761-006A-1 (1-3215) x US-10-209-264-2 (1-843)
QY 1 CTCACAAATCCACCAAGCTCTGTAGATCCAGGGTGAGGGGCTATATTTCTCTGC 60
DB 304 LeuHisAsnIleProSerSerAlaArgSerGlnGlyGluGlyProIlePheSerCys 323
QY 61 TGTGTGCTCCAGTCCGACAGTAACCTGTTCGACTACTCTCTCCATATCGTC 120
DB 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerHisIleVal 343
QY 121 AATCTTCTCGAGACTGGGACCCCTGCACGAACTGGAGAACACAACTCAGGATTCCT 180
DB 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
QY 181 AGGACCCCTGCTGCTTTACAGCGGGTTTTCTCGTTGACAAAGAACTCTCACAAATACC 240
DB 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
QY 241 GCAGACTAGACTCTGTGGACTTCTCAATTTCTAGGGGGAGCACCGGTTCCTCC 300
DB 384 AlaGluSerArgLeuTrpTrpThrSerLeuAsnPheLeuGlyAlaProThrCysSer 403
QY 301 TGGCCAAATTCGAGTCCCACTCCAACTCCTACCACTCCTGTCCTCCCAATTG 360
DB 404 TrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeu 423
QY 361 TCTGTGCTATCGCTGGATGTCTGCGGGGTTTTATCATATTCCTTTCATCTGCTGCT 420
DB 424 SerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAla 443
QY 421 ATGCCTCATCTTCTGTGTGTTCTTCTGGACTACCAAGGTATGTTCCTGCTGCTCT 480
DB 444 MetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSer 463
QY 481 ACTTCCAGGAACATCAACACAGCAGCGGGCCCTCAAGACTGCAGCCTCTCTGCTCA 540
DB 464 ThrSerArgAsnIleAsnHisGlnHisGlyAlaMetGlnAspLeuHisAspSerCysSer 483

```

```

QY 541 AGGAACTCTACGTTTCCCTTGTCTGTACAAAACCTTCGACGGAACCTGCACCTTG 600
DB 484 ArgLysLeuTyrValSerLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLeu 503
QY 601 TATTTCCCATCCATCATCTCGGCTTTTCGCAAGATTTCTATGGAGTGGGCTCAGTCCG 660
DB 504 TyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPro 523
QY 661 TTTCTCTGGCTCAGTTTACTAGTGCATTTGTTTTCAGTGGTTCGTAGGGCTTTCCCCAC 720
DB 524 PheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHis 543
QY 721 TGTGTGCTTTCAGTTATATGATGATGTGTATTTGGGGCGGAAGTCTGTACACATCTT 780
DB 544 CysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLeu 563
QY 781 GAGTCCCTTTTACTCTTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCT 840
DB 564 GluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPro 583
QY 841 AATAAAACCAACGTTGGGCTACTCCCTTAACCTTCATGGGATATGTAATTTGGAAGTTGG 900
DB 584 AsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTrp 603
QY 901 GGTACTTTACCGCAGGACATATTGTACTAAACTCAAGCAATGTTTTCGAAATCGCT 960
DB 604 GlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPro 623
QY 961 GTAAATAGACTATTGATTGAAAAGTATGTCAAAGAATTTGGGGTCTTTTGGGCTTTGCT 1020
DB 624 ValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAla 643
QY 1021 GCGCTTTTACCAATGTGGCTATCTCTGCTGCTGATGCTTTTATATGATGATATCAATCT 1080
DB 644 AlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSer 663
QY 1081 AAGCGGCTTTCACCTTTCTCCCAACTTACAGGCTTTCTGTGTAAACATATCTCGAAC 1140
DB 664 LysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAsn 683
QY 1141 CTTTACCCTGTTGCCGCAACCGTCCGCTCTCTGCCAAGTGTTCGTGACGCAACCCCT 1200
DB 684 LeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPro 703
QY 1201 ACTGGATGGGCTTGGCCATAGGCCATCAGCGCATGCTGGAACTTCTTGCTCTCTG 1260
DB 704 ThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetAlaGlyThrPheLeuAlaProLeu 723
QY 1261 CGATCCATCTAGCGGAACCTCTAGCAGCTTGTTCGTCGAGCGGCTCTGGAGCAAAA 1320
DB 724 ProIleHisThrAlaGluLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys 743
QY 1321 CTTATCGGAACCGCAACTCTGTTGCTCTCTCGAAATACACCTCTCTTCATGGCTG 1380
DB 744 LeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProIlePro 763
QY 1381 CTAGGCTGTGTCGCAACTGGATCTGCGCGGAGCTCTTTTGTCTACGTCCTCCGTCGG 1440
DB 764 LeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAla 783
QY 1441 CTGAATCCCGGAGACGACCGCTCTCGGGCGGTTGGGGCTCTACGTCCTCTTCTCAT 1500
DB 784 LeuAsnProAlaAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHis 803
QY 1501 CTGCGGCTTCGCGCGGACCGGCGGACCTCTCTTTACGGGCTCTCCCGTATGCT 1560
DB 804 LeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValLysProIlePro 823
QY 1561 TCTCATCTGCGGAGCGGTGTGCTTCTGCTTCTGCTTCTGCTGCTGCTGCTGCTGCT 1620
DB 824 SerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaIleTrpArgProPro 843

```

RESULT 3

US-08-591-502B-45
; Sequence 45, Application US/08591502B
; Patent No. 6607727
; GENERAL INFORMATION:
; APPLICANT: Chisari, Francis V.
; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
; Lymphocyte Responses to Hepatitis B Virus
; NUMBER OF SEQUENCES: 99
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,502B
; FILING DATE: 20-May-1996
; CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 08/100,870
; FILING DATE: 02-AUG-1993
; APPLICATION NUMBER: WO PCT/US94/08685
; FILING DATE: 01-AUG-1994

ATTORNEY/AGENT INFORMATION:

; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 014740-000230US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 45:

SEQUENCE CHARACTERISTICS:

; LENGTH: 843 amino acids

; TYPE: amino acid

; STRANDEDNESS: <Unknown>

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 45:

US-08-591-502B-45

Alignment Scores:

Pred. No.:	5e-248	Length:	843
Score:	2807.00	Matches:	528
Percent Similarity:	98.33%	Conservative:	3
Best Local Similarity:	97.78%	Mismatches:	8
Query Match:	47.52%	Indels:	2
DB:	2	Gaps:	0

US-10-761-006A-1 (1-3215) x US-08-591-502B-45 (1-843)

QY	4	CACAAATCCACCAAGCTGCTAGATCCAGGGTGAGGGCCCTATATTTCTCTGCTGG	63
DB	305	HisAenlleProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCysTrp	324
QY	64	TGGCTCCAGTTCCGGAACAGTAACCTGTTCCGACTACTGCTCTCCCATATGTCAT	123
DB	325	TrpLeuGlnPheArgAsnSerIlyProCysSerAspTyrCysLeuThrHisIleValasn	344
QY	124	CTTCTCGAGGACTGGGGACCTCGACCGAACATATGGAGAACCAACATCAGGATTCCTAGG	183
DB	345	LeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIleProArg	364

QY	184	ACCCCTGCTGTTTACAGGGGGGGTCTTCTCGTTGACAAGAATCTCTCAATACCGCA	243
DB	365	ThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThrThr	384
QY	244	GAGTCTAGACTCTG-GTGGACTTCTCTCAATTTTCTAGGGGGAGCACCCAGTTCCTG	302
DB	385	GluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-SerTr	404
QY	303	GCACAAATTCGCACTCCCAACCTCCAATCACCACTCTTGTCTCTCCAATTTGTC	362
DB	404	pProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeuSe	424
QY	363	CTGGCTATCGCTGATGTCCTGGCGGTTTATCATATTCCTCTTCATCTCTGCTAT	422
DB	424	rTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAlaMe	444
QY	423	GCCTCATCTTCTTGGTCTCTCGACTACCAAGTATGTGCCCCGTTCCTCTCTAC	482
DB	444	tProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSerTh	464
QY	483	TTCCAGGAACATCAACCAACAGCAGCGGGGCATGCAAGACTGCACCACTCTCTCAAG	542
DB	464	rSerArgAsnIleasnHisGlnHisGlyAlaMetGlnAspLeuHisaspSerCysSerAr	484
QY	543	GAACCTCTACGTTTCCCTCTGTTGCTGTACAAACCTTCGGACGGAACATGCACTTGT	602
DB	484	gAsnLeuTy-ValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLeuTy	504
QY	603	TTCCCATCCCATCATCTGGGCTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCGTT	662
DB	504	rSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerProPh	524
QY	663	TCTCTGCTCAGTTTACTAGTGCATTTGTTTCAGTGGTTCGTAGGGCTTTCCCCACTG	722
DB	524	eLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHisCy	544
QY	723	TTTGGCTTTTCAGTTATATGATGATGTTGTTATTTGGGGCGAAGTCTGTACAACATCTGA	782
DB	544	sLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLeuGl	564
QY	783	GTCCCTTTTACCTCTATTACCAATTTCTTTGCTCTTGGGTATATATTAAACCTTAA	842
DB	564	uSerLeuPheThrSerIleThrAsnPheLeuLeuLeuSerLeuGlyIleHisLeuAsnProAs	584
QY	843	TAAACCAAAACGTTGGGGCTACTCCCTTAACTTCATGGGATATGTAATTTGGAAGTTCGGG	902
DB	584	nLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTrpGl	604
QY	903	TACTTTACCGAGAACATATTTGACTTAAACCTCAAGCAATGTTTTCGAAAACCTGCTGT	962
DB	604	yThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuProVa	624
QY	963	AAATAGACCTTATTCATTTGGAAGTATGTCAAAGTATGTCGGTCTTTTGGGCTTGTCTGC	1022
DB	624	lAsnSerProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAlaAl	644
QY	1023	CCCTTTTACAAATGTGGCTATCTCGCTTGTAGCGCTTTATATGTCATGTATACAATCTAA	1082
DB	644	aProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSerLy	664
QY	1083	GCAGGCTTTTCATTTTCTCGCAACTTACAAGGCTTTCTGTGTAAACAATATCTGAACCT	1142
DB	664	sGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAsnLe	684
QY	1143	TTACCCCGTTCGCGGCAACGCTCGCTCTCGCAAGTCTTTCGCAAGTCTTTCGCAAGCCCCAC	1202
DB	684	uTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrProTh	704
QY	1203	TGGATGGGGCTTGGCCATAGCCATCAGCGCATGGCTGGGAACCTTCTTCTGCTCTCTGCC	1262
DB	704	rGlyTrpGlyLeuAlaIleGlyHisArgA-gMetArgGlyThrPheValAlaProLeuPr	724
QY	1263	GATCATCTACTCGGGAACCTCTAGACGCTTGTTTTGTTCGCAGCGCGGTCTGGAGCAAACT	1322

Db 724 oileHistrAlaGluLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaAlaYsLe 744
QY 1323 TATCGAACCGCAACTGTGTCTCTCTCGGAAATACACTCTCTTTCATGCTGT 1382
Db 744 uileGlyThrAspAsnSerValValSerArgLysTyThrSerPheProTrpLeuLe 764
QY 1383 AGGTTGTCTGCAACTGATCCTCGCGGAGACGTCTTTGTCTACGTCCGTCGCGCT 1442
Db 764 uGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyValProSerAlaLe 784
QY 1443 GAATCCCGGAGCGCGCTCTCGGGCGGTGTGGGCTCTACCGTCCCTCTCTCACT 1502
Db 784 uAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyArgProLeuLeuLe 804
QY 1503 GCCGTTCCGCGCGCACACGCGGCGCACCTCTCTTACGCGGTCTCCCGCTATGTGCTTC 1562
Db 804 uProPheArgProThrThrGlyArgThrSerLeuTyAlaValSerProSerValProSe 824
QY 1563 TCATTCGCGGACCGGTGTCACCTTCGTCACCTCTGCACTGCGATGAGACCGCG 1620
Db 824 rHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPro 843

RESULT 4

US-08-591-502B-59
; Sequence 59, Application US/08591502B
; Patent No. 6607727
; GENERAL INFORMATION:
; APPLICANT: Chieari, Francis V.
; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
; Lymphocyte Responses to Hepatitis B Virus
; NUMBER OF SEQUENCES: 99
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,502B
; FILING DATE: 20-May-1996
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 08/100,870
; FILING DATE: 02-AUG-1993
; APPLICATION NUMBER: WO PCT/US94/08685
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 014740-0002300S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 843 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 59:
US-08-591-502B-59

Alignment Scores: 5e-248 Length: 843
Pred. No.: 2807.00 Matches: 527
Score: 98.34% Conservative: 5
Best Local Similarity: 97.41% Mismatches: 8
Query Match: 47.52% Indels: 2
Gaps: 0
US-10-761-006A-1 (1-3215) x US-08-591-502B-59 (1-843)
QY 1 CTCCACAACTTCACCAAGCTCTGTAGATCCAGGTCAGGGGCTATATTTTCCTGC 60
Db 304 LeuHisAsnIleProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCy 323
QY 61 TGGTGCTCCAGTTCGGAACAGTAAACCTGTTCGACTACTGCTCTCCATATGCTC 120
Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrcysLeuThrHisIleVal 343
QY 121 AATCTTCTGAGGACTGGGGACCTTGACCGAACATGAGAGACACAACTCAGGATTCT 180
Db 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
QY 181 AGGACCCCTGCTGTTTACAGCGCGGTTCCTGTTGACAAAGATCCTCACAAATACC 240
Db 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAnThr 383
QY 241 GCAGAGTCTAGACTCTG-GTGGACTTCTCTCAATTTCTAGGGGAGACCCACGTGTC 299
Db 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
QY 300 CTGGCCAAATTCGAGTCCCACTCAATCACTCACTCACTCACTCACTCTGTCTCCAAATTT 359
Db 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
QY 360 GTCTCTGCTATCGCTGGATGTCTGCGCGCTTTTATCATATTCCTCTTCCTCTGCTGC 419
Db 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyHisIleProLeuHisIleProAla 443
QY 420 TATGCTCATCTTCTTGTGTTCTCTGGAATACCAAGGTATGTTGCCGCTGTGCTC 479
Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyValAlaArgLeuSerSe 463
QY 480 TACTTCCAGGAACATCAACACGACGAGCGGCGCATGAGACCTGACGACTCTCTGCTC 539
Db 463 rThrSerArgAsnIleAsnTyrglnHisGlyThrMetGlnAsnLeuHisAspSerCysSe 483
QY 540 RAGGAACCTCTACGTTTCCCTCTGTTGCTGTACAAACCTTCGGACGGAACCTGACACTT 599
Db 483 rArgAsnLeuTyValSerLeuLeuLeuLeuTyHisThrPheGlyArgLysLeuHisLe 503
QY 600 GTATTCCCATCCCATCATCTCTGGGCTTTTCGCAAGATTCCTATGGGAGTGGGCTCAGTCC 659
Db 503 uTySerHisProIleLeuLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
QY 660 GTTCTCTCGGCTCAGTTTACTAGTGCATTTGTTGTCAGTGGTTCGTAGGGCTTCCCCCA 719
Db 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 543
QY 720 CTGTTTGGCTTTTCAGTTATATGATGATGATGATGATGATGATGATGATGATGATGAT 779
Db 543 sCysLeuAlaPheSerTyMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTTGGGTATACATTTAAACCC 839
Db 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAnPr 583
QY 840 TAATAAAACCAACCGTTGGGCTACTCCCTTAATTTTATGGATATGTAATTCGAAGTTG 899
Db 583 oAsnLysThrLysArgTrpGlyTy-SerLeuAsnPheMetGlyTyrValIleGlySerTr 603
QY 900 GGGTACTTTTACCGCAGGAACATATGTTACTTAAACTCAAGCATGTTTTCGAAACTGCC 959

```

Db 603 pGlyThrLeuProGlnGluHisIleValGlnLysLeuLysGlnCysPheArgLysLeuPr 623
Qy 960 TGTAATAGACCTATGTGTTGAAAGATGTCACAAAGATTTGGGTCTTTTGGGGCTTTC 1019
Db 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
Qy 1020 TGCCCTTTTACACATGTCCTATCTCCCTGCTTATATGCTTATATGCTATATATC 1079
Db 643 aAlaProPheThrGlnCysGlyTrpProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
Qy 1080 TAAGCAGGCTTTCACTTTCTGCCCACTTACAGGCTTCTGTGTTAAACAATATCTGAA 1139
Db 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysGlnTyrLeuAs 683
Qy 1140 CTTTACCCCTGTCGCGCAACGTCGGTCTCTGCCAAGTGTGTTGTCGACGCAACCCC 1199
Db 683 nLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
Qy 1200 CACTGGATGGGCTTGGCCATAGGCCATCAGGCATGCTGGAACTTCTTCTGCTCTCT 1259
Db 703 oThrGlyTrpGlyLeuAlaIleGlyHisArgMetArgGlyThrPheValAlaProLe 723
Qy 1260 GCCGATCCATCTAGCGGAACCTCAGCAGCTTGTTCCTCGCAGCGCTCTGGACAA 1319
Db 723 uProlleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLy 743
Qy 1320 ACTTATCGGAACCGCAACTGTGTGCTCTCTCGGAAATACACCTCTTTCATGGCT 1379
Db 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 763
Qy 1380 GCTAGGTTGCTGCAACTGATCTCGCGGAGAGTCTTGTCTACGTCCCGTGGC 1439
Db 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
Qy 1440 GCTGAATCCCGGAGACCGCTCTCGGGGCGCTTGGGGCTCTACGCTCCCTTCTCA 1499
Db 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuH 803
Qy 1500 TCTCGCTTCCGGCGCAGCACCGGCGCACCTCTCTTTACCGGTCTCTCCCGTATGTGCC 1559
Db 803 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
Qy 1560 TTCTCATCTCGCGGACCGTGTGCACTTCGCTTACCTTGCACGTGCGATGAGACACCC 1619
Db 823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 843
Qy 1620 G 1620
Db 843 o 843

```

RESULT 5

```

US-08-591-502B-47
; Sequence 47, Application US/08591502B
; Patent No. 6607727
; GENERAL INFORMATION:
; APPLICANT: Chisari, Francis V.
; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
; Lymphocyte Responses to Hepatitis B Virus
; NUMBER OF SEQUENCES: 99
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,502B

```

```

; FILING DATE: 20-May-1996
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 08/100,870
; FILING DATE: 02-AUG-1993
; APPLICATION NUMBER: WO PCT/US94/08685
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 014740-000230US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 843 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 47:
US-08-591-502B-47

Alignment Scores:
Pred. No.: 2,75e-246 Length: 843
Score: 2788.00 Matches: 523
Percent Similarity: 97.97% Conservative: 7
Best Local Similarity: 96.67% Mismatches: 10
Query Match: 47.20% Indels: 2
DB: 2 Gaps: 0

US-10-761-006A-1 (1-3215) x US-08-591-502B-47 (1-843)
Qy 1 CTCACAACTTCACCAAGCTCTGCTAGATCCAGGGTGGGGCCCTATATTTCTCTGC 60
Db 304 LeuHisAsnIleProProSerSerAlaArgSerGlnSerGlnGlyProIlePheSerCys 323
Qy 61 TGTGGTCTCCAGTTCGGGAACAGTAAACCTGTTCGAGTACTGCTCTCCCATATCGTC 120
Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 343
Qy 121 AATCTTCTCGAGGACTGGGACCTCGACCGACATCGAGACACACACATCAGGATTCCT 180
Db 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGluHisAsnIleArgIlePro 363
Qy 181 AGGACCCCTGCTGCTGTATTACAGCGGGGTTTTCTCGTTGACAGAATCTCTCAATATACC 240
Db 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
Qy 241 GCAGAGTCTAGACTCTG-GTGGACTTCTCTCAATTTTCTAGGGGAGACCCACGTGTC 299
Db 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
Qy 300 CTGCGCCAAATTCGACGTCCCACTCCCAATCACTCACCACCTCTGTGCTCTCAATTT 359
Db 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
Qy 360 GTCTGCTATCGCTGATGTGTCGCGGGGTTTTATCATATCTCTTCTATCTCTGCTGC 419
Db 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
Qy 420 TATGCTCATCTTCTTCTGTTCTTCTGACTACCAAGGTATGTTCCCGCTTGTCTCTC 479
Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 463
Qy 480 TACTTCCAGGACATCAACACAGCAGCGGGCCATCGACAGCTCTGACCTCTCTCTCTC 539
Db 463 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483

```

```

QY 540 AAGGAACTCTACGTTTCCCTTGTGCTGTACAAAACCTTCGACGGAACCTGCACCTT 599
Db |||||
QY 483 rArgAsnLeuYrValSerLeuLeuLeuYrLysThrPheGlyArgLysLeuHisLe 503
Db |||||
QY 600 GTATTCCCATCCATCATCTGGGCTTTCGCAAGATTCTATGGAGTGGGCTCAGTCC 659
Db |||||
QY 503 uYrSerHisProIleIleuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
Db |||||
QY 660 GTTCTCTCGGCTCAGTTTACTAGTCCCAATTTGTCAGTGGTTCGTAGGGCTTTCCCA 719
Db |||||
QY 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 543
Db |||||
QY 720 CTGTTTGGCTTTCAGTATATATGATGATGTGTATTTGGGGCGAAGTCTGTACAATCT 779
Db |||||
QY 543 sCysLeuAlaPheSerTyMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
Db |||||
QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCC 839
Db |||||
QY 563 uGluSerLeuPheThrSerIleThrAsnPheLeuSerLeuGlyIleHisLeuAsnPr 583
Db |||||
QY 840 TAATAAAACCAACGTTGGGCTACTCCCTTAACCTTCATGGGATATGTAATTGGAAGTTG 899
Db |||||
QY 583 oAsnLysThrLysArgTrpGlyTySerLeuAsnPheMetGlyTyValIleGlySerTr 603
Db |||||
QY 900 GGGTACTTTACCGCAGGAACATATGTTACTATAAACTCAAGCAATGTTTTCGAAAACCTGCC 959
Db |||||
QY 603 pGlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPr 623
Db |||||
QY 960 TGTAAATAGACTATTGATTGGAAGTANGTCAAGAATTGGGTCTTTTGGCTTTCGCTTGC 1019
Db |||||
QY 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
Db |||||
QY 1020 TGCCCTTTTACACAAATGCGGTATCTCCTGCTTATGTCCTTATATGATGATATCAATC 1079
Db |||||
QY 643 aAlaProPheThrGlnCysGlyTyProAlaLeuMetProLeuTyAlaCysIleGlnSe 663
Db |||||
QY 1080 TAAGAGGCTTTCACCTTCTCGCAACTTCAAGGCTTCTGTGTAAACAATATCTGAA 1139
Db |||||
QY 663 rLysGlnAlaPheThrPheSerProThrTyLysAlaPheLeuCysGlnGlnTyLeuHi 683
Db |||||
QY 1140 CTTTACCCGTTGCGCGCAAGGTCGGTCTCTCCCAAGTGTGTTGTCGACGCAACCC 1199
Db |||||
QY 683 sLeuTyProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
Db |||||
QY 1200 CACTGATGGGCTTGGCCATAGCCATCAGCGCTGCTGGAACCTTTCTGCTCTCTCT 1259
Db |||||
QY 703 oThrGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValValProLe 723
Db |||||
QY 1260 GCCGATCCATACCTGCGGAACCTCTAGCAGCTTGTGTTGTCGAGCGGCTCGGAGCAA 1319
Db |||||
QY 723 uProIleHisThrAlaGluLeuAlaCysPheAlaArgAspArgSerGlyAlaLys 743
Db |||||
QY 1320 ACTTATCGGAACCGACACTCTGTTCTCTCTCGGAATACACCTCTTCCATGGCT 1379
Db |||||
QY 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyThrSerPheProTrpLe 763
Db |||||
QY 1380 GCTAGGCTGTGTCGCAACTGGAATCTGCGCGGAGCTCTTGTGTCTAGCTCCCGTCGGC 1439
Db |||||
QY 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyValProSerAl 783
Db |||||
QY 1440 GCTGAATCCCGCGGACGACCGCTCTCGGGCCGTTTGGGGCTCTACCGTCCCTTCTTCA 1499
Db |||||
QY 783 aLeuAsnProAlaAspAspProSerArgLysGlyLeuGlyLeuTyArgProLeuLeuSe 803
Db |||||
QY 1500 TCTGCGGTTCCGCGGACGACGCGGCGACCTCTCTTACGCGGCTCTCCCGTATGTGCC 1559
Db |||||
QY 803 rLeuProPheGlnProThrThrGlyArgThrSerLeuTyAlaValSerProSerValPr 823
Db |||||
QY 1560 TTCTCATCTCGCGGACCGTGTGACTTCTGCTTCTCACTCTGACGTCGTCATGGAGACACC 1619
Db |||||
QY 823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 843
Db |||||

```

```

QY 1620 G 1620
Db 843 o 843

RESULT 6
; US-08-591-502B-46
; Sequence 46, Application US/08591502B
; Patent No. 6607727
; GENERAL INFORMATION:
; APPLICANT: Chisari, Francis V.
; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
; Lymphocyte Responses to Hepatitis B Virus
; NUMBER OF SEQUENCES: 99
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,502B
; FILING DATE: 20-May-1996
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 08/100,870
; FILING DATE: 02-AUG-1993
; APPLICATION NUMBER: WO PCT/US94/08685
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 014740-000230US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 845 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 16..18
; OTHER INFORMATION: /product= "OTHER"
; /note= "Xaa = unknown"
; SEQUENCE DESCRIPTION: SEQ ID NO: 46:
US-08-591-502B-46

Alignment Scores:
Pred. No.: 6,528-245 Length: 845
Score: 2773.00 Matches: 519
Percent Similarity: 97.04% Conservative: 6
Best Local Similarity: 95.93% Mismatches: 15
Query Match: 46.94% Indels: 2
DB: 2 Gaps: 0

US-10-761-006A-1 (1-3215) x US-08-591-502B-46 (1-845)

QY 1 CTTCCACAACTTCCCAAGCTCTGCTAGATCCAGGTGAGGGCTATATTTCTCTGC 60
Db |||||
QY 306 LeuHisAsnIleProProSerSerAlaArgProGlnSerGlyGlyProIleLeuSerCys 325
Db |||||

```

QY 61 TGTGGCTCCAGTTCCTCGGAAACAGTAACCCCTGTTCGGAAGTCTGCTCCATATCCTC 120
Db 326 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 345
QY 121 AATCTTCTCGAGGACTGGGACCTCGACCAACATGAGAGACACACATCAGATTCCT 180
Db 346 AsnLeuLeuGluAAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 365
QY 181 AGGACCCCTGCTCGTGTACAGGGGGGTTTTCTCGTTTGACAAAGATCCTCACAAATACC 240
Db 366 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 385
QY 241 GCAGAGTCTAGACTCTG-GTGAGACTTCTCAATTTCTAGGGGAGACCCACGTGTTC 299
Db 386 ThrGluSerThrLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 405
QY 300 CTGGCCAAATTCGAGTCCCAACCTCCATCACTACCAACCTCTGTCTCTCAATTT 359
Db 405 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 425
QY 360 GTCTGGCTATCGTGGATGTCTGCGGCGTTTTATCATATTCCTCTTCATCCTGCTGC 419
Db 425 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 445
QY 420 TATGCTCATCTCTGTGTGTTCTCTGGAAGTATGTTGCGGCTTTGTCCTC 479
Db 445 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValValCysLeuSerSe 465
QY 480 TACTTCCAGGAACATCAACACCCAGCAGCGGGCCATGCAAGCTGACGACTCTGCTC 539
Db 465 rThrSerLysAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 485
QY 540 AAGGAACTTACGTTTCCCTCTGTTGCTGTACAAACCTTCGAGCAACCTGCACTT 599
Db 485 rArgAsnLeuTyrValSerLeuPheLeuLeuTyrHisThrPheGlyArgLysLeuHisLe 505
QY 600 GTATTCCCATCCATCATCTCGGGCTTCGCAAGATTCCTTAGGAGTGGCCCTCAGTCC 659
Db 505 uTyrSerHisProIleLeuLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 525
QY 660 GTTTCCTCGCTCAGTTTACTAGTGCATTTGTTTCAGTGGTTCGTAGGGCTTTCCCCA 719
Db 525 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi 545
QY 720 CTGTTTGGCTTTCAGTTATATGATGTGGTATTGGGGCGAAGTCTGTACAACTCT 779
Db 545 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 565
QY 780 TGAGTCCCTTTTACCTCTATTACAAATTTCTTTTGTCTTTGGGTATACATTTAAACCC 839
Db 565 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 585
QY 840 TAATAAACCAACAGTTGGGCTACTCCCTTAACCTTCATGGGATGTAAATGGAAGTTG 899
Db 585 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlyCysTr 605
QY 900 GGGTACTTTACCGAGGAACATATTGTACTAAACTCAAGCAATGTTTTTCGAAACTGCC 959
Db 605 pGlyThrLeuProGlnGluHisIleValLeuLysIleLysGlnCysPheArgLysLeuPr 625
QY 960 TGTAAATAGACTATTTGATGGAAGTATGTCAAGAAATTCGGGTCTTTTGGGCTTTCG 1019
Db 625 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 645
QY 1020 TGCCCTTTTACAAATGTGGCTCTCGCTTCATGCTTTTATGTCATGTATACATC 1079
Db 645 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 665
QY 1080 TAAGCAGGCTTTCATCTTCTCGCCCAACTTCAAGGCTTCTCTGTGTAACCAATATCTGAA 1139
Db 665 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 685

QY 1140 CCTTTACCCCTTCCCGGCAACCGTCCGFTCTCTGCCAAGTGTGTTGCTGACGCAACCCC 1199
Db 685 nLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 705
QY 1200 CACTGGATGGGCTTGCCCATAGCCATCAGCGCATGGCTGGAACTTCTTGGCTCCTCT 1259
Db 705 oThrGlyTrpGlyLeuAlaIleGlyHisArgMetArgGlyThrPheValAlaProLe 725
QY 1260 GCCCATCCATACTCCGCAACTCTCAGCAGCTGTGTTTGTCTCGCAGCCGCTCTGGAGCAA 1319
Db 725 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys 745
QY 1320 ACTTATCGGAACCAACTCTGTGTCTCTCTCGGAATAACACCTCTTCCATGGCT 1379
Db 745 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 765
QY 1380 GCTAGGTGTGCTCCCAACTCGATCTCGGGGCGCTCTTCTTCTACGTCGCTCGGC 1439
Db 765 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 785
QY 1440 GCTCAATCCCGCAGCAGCCGCTCTCGGGGCGCTTGGGGCTCTACCGTCCCTTCTTCA 1499
Db 785 aLeuAsnProAlaAspPProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 805
QY 1500 TCTCCGCTTCCGCGCAGCAGCGGCGCACCTCTCTTTACGCGTCTCCCGTATGTGCC 1559
Db 805 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 825
QY 1560 TTCTCATCTCCGCGCAGCTGTGCACTTCGTTTCACTCTCGACGTCGACGTCGAGACACC 1619
Db 825 oSerHisLeuProAspArgValHisPheProSerProLeuHisValAlaTrpArgProPr 845
QY 1620 G 1620
Db 845 o 845
RESULT 7
US-08-591-502B-48
; Sequence 48, Application US/08591502B
; Patent No. 6607727
; GENERAL INFORMATION:
; APPLICANT: Chisari, Francis V.
; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
; Lymphocyte Responses to Hepatitis B Virus
; NUMBER OF SEQUENCES: 99
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,502B
; FILING DATE: 20-May-1996
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 08/100,870
; FILING DATE: 02-AUG-1993
; APPLICATION NUMBER: WO PCT/US94/08685
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762

; REFERENCE/DOCKET NUMBER: 014740-000230US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 576-0200

; TELEFAX: (415) 576-0300

; INFORMATION FOR SEQ ID NO: 48:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 843 amino acids

; TYPE: amino acid

; STRANDEDNESS: <Unknown>

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 48:

US-08-591-502B-48

Alignment Scores:

Pred. No.:	1.25e-243	Length:	843
Score:	2759.00	Matches:	519
Percent Similarity:	97.23%	Conservative:	7
Best Local Similarity:	95.93%	Mismatches:	14
Query Match:	46.71%	Indels:	2
DB:	2	Gaps:	0

US-10-761-006A-1 (1-3215) x US-08-591-502B-48 (1-843)

```
QY 1 CTCACACATTCACCAAGCTCTGCTAGATCCCAAGGTGAGGGCTATATTTTCCTGC 60
DB 304 LeuHisAsnIleProProSerCysAlaArgSerGlnSerGluGlyProIleSerSerCys 323
QY 61 TGGTGGCTCCAGTTCGGGAACAGTAACCCCTGTTCCGACTACTGCTCTCCATATCGTC 120
DB 324 TrpIleGlnPheArgAsnSerGluProCysSerAspTyrCysLeuThrHisIleVal 343
QY 121 AATCTTCTCGAGGACTGGGGACCCCTGCACCCGAACATGGAGAACACCAATCAGGATTCCT 180
DB 344 AsnLeuLeuGluAspTyrGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
QY 181 AGGACCCCTGCTGTTACAGCGGGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
DB 364 ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 383
QY 241 CGAGAGTCTAGACTCTGCTCAATTTCTAGGGGAGCACCCACCTGCTTC 299
DB 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgIleSerThrHisVal-Se 403
QY 300 CTGGCCAAATTCGAGTCCCCCAACTCCAATCACTCAACCACTCTTGTCTCCCAATTT 359
DB 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuSerSerAsnLe 423
QY 360 GTCTGGCTATCGCTGGATGTCCTGGCGGCTTTTATCATATTCCTCTTCATCTCTGCTGC 419
DB 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
QY 420 TATGCTCATCTCTTGTGTTCTTCGACTTACCAAGTATGTTGCCGTTGCTCTCTC 479
DB 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 463
QY 480 TACTTCCAGGAACATCAACACCAAGCACGGGGCCATGCAAGACCTGCAGCACTCTCTGCTC 539
DB 463 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
QY 540 AAGGAACCTCTAGTTTCCTCTGTTGCTGTGTATACAAAACCTTCGGACGGAACCTGACTT 599
DB 483 rArgAsnLeuTyrValSerLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 503
QY 600 GTATTTCCCATCCATCATCTGGGCTTCGCAAGATTCCTATGGAGTGGGCTCAGTCC 659
DB 503 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
QY 660 GTTTCCTCGGCTCAGTTACTAGTGCATTTGTTTCAGTGGTTCGTAGGGCTTTCCCCCA 719
DB 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi 543
QY 720 CTGTTTGGCTTTCAGTTATATGATGATGTTGTTGGGGGCGAAGTCTGTACAACATCT 779
```

```
DB 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
QY 780 TGAATCCCTTTTACCTCTATACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCC 839
DB 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
QY 840 TAATAAAACCAAGCTTGGGGCTACTCCCTTAACCTTCATGGGATATGTAATTGGAAGTTG 899
DB 583 oHisLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
QY 900 GGGTACTTTTACCGCAGGAACATATTGTACTAAAACTCAAGCAATGTTTTCGAAAACCTGCC 959
DB 603 pGlyThrLeuProGlnGluHisIleValLeuLysIleLysGlnCysPheArgLysLeuPr 623
QY 960 TGTAAATAGACCTATTGATTTGAAAGTATGATCAAGAAATTTGTGGGTCTTTTGGGGCTTGC 1019
DB 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
QY 1020 TGCCCTTTTACACAATGTGGCTATCTGCTCTGCTGCTTTATATGATCATGTATACAATC 1079
DB 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
QY 1080 TAAGCAGGCTTTCTCTTCGCCAATCTTCAAGGCCCTTTCTGTGTAAACAATATCTGAA 1139
DB 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuHi 683
QY 1140 CTTTACCCCTTCCCGGCAACCGTCCGCTCTCTGCCAAGTGTGTTGCTGACGCAACCC 1199
DB 683 sLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
QY 1200 CACTGGATGGGGCTTGGCCATAGCCATCAGCGCATGGCTGGAACTTTCTGGCTCTCT 1259
DB 703 oThrGlyTrpGlyLeuAlaIleGlyGlnSerGlyMetArgGlyThrPheValAlaProLe 723
QY 1260 GCCGATCCATCTCGGAACTCTAGCAGCTTGTGTTTGTCTCGCAGCCGCTCTGGAGCAAA 1319
DB 723 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLy 743
QY 1320 ACTTATCGGAACCGACAACTCTGTGTCTCTCTCGGAAATACACCTCTTTCATGGCT 1379
DB 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrple 763
QY 1380 GCTAGGTGTCTGCTCCCACTGATCTCGCGGAGCGTCTTGTCTACGTCCCTCCGCGC 1439
DB 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
QY 1440 GCTGAATCCCGCGAGCAGCCGCTCTCGGGGCGGCTTGGGGCTCTACCTGCCCTTCTTCA 1499
DB 783 aLeuAsnProAlaAspAspProSerArgLysGlyArgLeuGlyLeuTyrArgProLeuLeuHi 803
QY 1500 TCTCGCTTCGGCGCGACCCAGCGGCGCACCTCTCTTTACGGGCTCTCCCGTATGTGCC 1559
DB 803 sLeuProPheArgProThrThrGlyArgAlaSerLeuTyrAlaValSerProSerValPr 823
QY 1560 TTCTCATCTCGGAGCCGTGTGCATCTGCTTCACTCTGCACTGCGCATGCGATGAGACACC 1619
DB 823 cSerHisLeuProValArgValHisPheAlaSerProLeuHisValAlaIleTrpArgProPr 843
QY 1620 G 1620
DB 843 o 843
```

RESULT 8

US-08-416-950-11

; Sequence 11, Application US/08416950

; Patent No. 5780036

; GENERAL INFORMATION:

; APPLICANT: CHISARI, Francis V.

; TITLE OF INVENTION: PEPTIDES FOR INDUCING CYTOTOXIC T

; TITLE OF INVENTION: LYMPHOCYTE RESPONSES TO HEPATITIS B VIRUS

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: One Market Plaza, Steuart Street Tower
CITY: San Francisco
STATE: CA
COUNTRY: U.S.A.
ZIP: 94105-1492
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/416,950
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US
FILING DATE:
APPLICATION NUMBER: US 07/935,898
FILING DATE: 26-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,540
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: Parmelee, Steven W.
REGISTRATION NUMBER: 31,990
REFERENCE/DOCKET NUMBER: 14740-2-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 467-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 845 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-416-950-11

Alignment Scores:
Pred. No.: 1,25e-243 Length: 845
Score: 2759.00 Matches: 516
Percent Similarity: 97.23% Conservative: 10
Best Local Similarity: 95.38% Mismatches: 14
Query Match: 46.71% Indels: 2
DB: 1 Gaps: 0

US-10-761-006A-1 (1-3215) x US-08-416-950-11 (1-845)

```
Qy 1 CTCACAACTTCACCAAGCTCTGTAGATCCCGGGTGGGGCCCTATATTTCTCTGC 60
Db 306 LeuHisAsnIleProProAsnSerAlaArgSerGlnSerGluGlyProValPheSerCys 325
Qy 61 TGTGGCTCCAGTTCGGAGACAGTAACCCCTGTTCCGACTACTGCTCTCCATATCGTC 120
Db 326 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 345
Qy 121 AATCTTCTCCAGGACTCGGACCTCGACCGAACATCGGAGACACACATCAGATTCTT 180
Db 346 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 365
Qy 181 AGGACCCCTGCTCGTGTGTACAGGGGGGTTTTCTCGTTGACAGAAATCTTCAAAATACC 240
Db 366 ArgThrProIleArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 385
Qy 241 GCAGAGCTAGACTCTG-GTGGACTTCTCTCAATTTCTAGGGGGAGACCCACGTTTC 299
Db 386 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrArgVal-Se 405
Qy 300 CTGGCCAAATTCGACGTCCCACTCCCAATCACTACCAACCTCTGTGCTCCCAATTT 359
Db 405 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 425
```

```
Qy 360 GTCCTGGCTATCGCTCGAGTGTCTGGCGGTTTTATCATATTTCTCTTCTCATCTGCTGC 419
Db 425 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 445
Qy 420 TATGCTCATCTCTTGTGGTCTTCTTGACTACCAAGGTATGTGCGGCTTGTGCTCCTC 479
Db 445 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaArgLeuSerSe 465
Qy 480 TACTTCCAGGAACATCAACACACAGACGGGGCCATGCAAGACTGCACACCTCTCTCTC 539
Db 465 rAsnSerArgIleIleAsnTyrGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe 485
Qy 540 AAGGAAACTCTACGTTTCCCTTGTGTGTGTGTAACAAACCTTCGGACGGAAATGCACTT 599
Db 485 rArgAsnLeuTyrValSerLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 505
Qy 600 GTATTGCCATCCCATCATCTGGGCTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCC 659
Db 505 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 525
Qy 660 GTTTCTCTGGCTCAGTTTACTAGTGCATTTCTTCAGTGTGTGCTGAGGCTTTCCCCCA 719
Db 525 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHI 545
Qy 720 CTGTTTGGCTTTCAGTTTATATGATGATGTGTATTGGGGCGGAAAGTCTGTACAACATCT 779
Db 545 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 565
Qy 780 TGAGTCCCTTTTACCTCTATTAACCAATTTTCTTGTCTTTCGGGTATATATTAAACCC 839
Db 565 uGluSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 585
Qy 840 TAATAAAACCAAAAGTTGGGGCTACTCCCTTAACTTTCATGGGATATGTAATTGGAAGTGTG 899
Db 585 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 605
Qy 900 GGGTACTTTTACCGCAGGAACATATTGTACTAAAACTCAAGCAATGTTTTCGAAACATGCC 959
Db 605 pGlyThrIleProGlnGluHisIleValGlnLysIleLysGlnCysPheArgLysLeuPr 625
Qy 960 TGTAAATAGACCTATTGATTGGAAAGTATGTCTCAAGAAATGTGGGCTTTTGGCTTGC 1019
Db 625 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 645
Qy 1020 TGCCCTTTTACACAATGTGCTCTATCTGCTGCTTATATGCTCTTATATCATGTATACAAATC 1079
Db 645 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 665
Qy 1080 TAAGCAGGCTTTTCACTTTCTCGCCAACTTACAAGGCTTTTCTGTGTAAACAATATCTGAA 1139
Db 665 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 685
Qy 1140 CTTTACCCCTGTCGGCGCAACGGTCCGCTCTCTGCAAGTGTTCCTCAGCAACCC 1199
Db 685 nLeuTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 705
Qy 1200 CACTGGATGGGCTTGGCCATAGCCATAGCGCATGCTGGAACCTTTCTGGCTCTCTCT 1259
Db 705 oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValAlaProLe 725
Qy 1260 GCCGATCCATACCTCGGAACTCTTAGCAGCTGTTTGTTCGCGAGCCGGTCTGGAGCAAA 1319
Db 725 uProIleHisThrAlaGluLeuLeuAlaIaCysPheAlaArgSerArgSerGlyAlaLy 745
Qy 1320 ACTTATCGGAACCGACAACTCTGTTGCTCTCTCTCGGAAATACACCTCTCTTCCATGGCT 1379
Db 745 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 765
Qy 1380 GCTAGGCTGTGCTGCCAACTGGATCTGCGCGGAGCGTCTCTGTCTACGCTCCGCTCGGC 1439
Db 765 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 785
Qy 1440 GCTGAATCCCGCGGAGACCCGCTCTCGGGGCGGCTTTCGGGCTCTTACCGCTCCCTTCTTCA 1499
```



```
QY 1080 TAAGCAGCGCTTTCACCTTCTCCGCACTTACAGGCGCTTCTGTGTAAACAATATCTGAA 1139
Db 665 rLysGlnAlaPheThrPheSerProThrTyLysAlaPheLeuCyLysGlnTyLysLeuAs 685
QY 1140 CTTTACCGCTTGC CGGCAACGCTCGGCTCTGCGCAAGCTGTTCGACGCAACCC 1199
Db 685 nLeuTyProValAlaArgGlnAArgProGlyLeuCyGlnValPheAlaSerAlaThrPr 705
QY 1200 CACTGATGGGGCTTGGCCATAGGCGCATCAGCGATGGCTGGAACTTCTTGGCTCCTCT 1259
Db 705 oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValAlaProLe 725
QY 1260 GCCATCCACTCTCGGGAACCTCCTAGCAGCTGTGTTTCTCGCAGCGCGTCTGGAGCAA 1319
Db 725 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLy 745
QY 1320 ACTTATCGAACCGACCACTCTGTGTCCTCTCTCGGAATACACCTCTTCCATGGCT 1379
Db 745 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyThrSerPheProTrpLe 765
QY 1380 GCTAGGGTGTCTGCCAACTGGATCCTCGCGGGACGCTCTTGTCTACGTCCTCGCTGGC 1439
Db 765 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyValProSerAl 785
QY 1440 GCTGAATCCCGCGACGACCGCTCTCGGGCGCGTTTGGGGCTCTACCGTCCCTCTTCA 1499
Db 785 aLeuAsnProAlaAspAspProSerArgGlyArgGlyLeuTyArgProLeuLeuAr 805
QY 1500 TCTCGCGTTCGGCGGACCGACCGGGCGGCACCTCTCTTTACGCGCTCTCCCGTATGGCC 1559
Db 805 gLeuProPheArgProThrThrGlyArgThrSerLeuTyArgValSerProSerValPr 825
QY 1560 TTCTCATCTCGCGGACCGTGTGCACCTTCCGCTTACCTCTGCACGCTCGCATGGAGCACCC 1619
Db 825 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 845
QY 1620 G 1620
Db 845 o 845

RESULT 10
US-08-591-502B-11
; Sequence 11, Application US/08591502B
; Patent No. 6607727
; GENERAL INFORMATION:
; APPLICANT: Chiesari, Francis V.
; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
; Lymphocyte Responses to Hepatitis B Virus
; NUMBER OF SEQUENCES: 99
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,502B
; FILING DATE: 20-May-1996
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 08/100,870
; FILING DATE: 02-AUG-1993
; APPLICATION NUMBER: WO FCT/US94/08685
```

```
;
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 014740-000230US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 845 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1..845
; OTHER INFORMATION: /product= "OTHER"
; /note= "xaa = any amino acid
; (<50% consensus)"
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-08-591-502B-11
```

```
Alignment Scores: 5.56e-242 Length: 845
Pred. No.: 2741.00 Matches: 514
Score: 96.30% Conservative: 7
Percent Similarity: 95.01% Mismatches: 19
Best Local Similarity: 46.40% Indels: 2
Query Match: 2 Gaps: 0
DB:
```

US-10-761-006A-1 (1-3215) x US-08-591-502B-11 (1-845)

```
QY 1 CTCCACAACTTCACCAAGCTCTGTAGATCCAGGTCAGGGGCTATATTTCTCTGC 60
Db 306 LeuHisAsn***ProProAsnSerAlaArgSerGlnSerGluGlyProValPheSerCys 325
QY 61 TGGTGGCTCCAGTTCGGGAACAGTAAACCTCTGTTCCGACTACTGCTCTCCCATATCGTC 120
Db 326 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrcysLeu***HisIleVal 345
QY 121 AATCTTCTCGAGGACTGGGACCTCGACCCGAAACATGGAGAACACAAATCATGAGATTCT 180
Db 346 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHis***IleArgGilePro 365
QY 181 AGACCCCTCTCGTGTACAGCGGGGTTTCTCGTTGACAGATCTCTCAATATCTCAATACC 240
Db 366 ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 385
QY 241 GCAGAGTCTAGACTCTG-GTGGACTTCTCTCAATTTTCTAGGGGAGACCCACGTCGTC 299
Db 386 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGly***ThrArgVal-Se 405
QY 300 CTGGCCAAAATTCGAGTCCCCAACCTCCAATCATCACTCACCAACCTCTGTCTCTCAATTT 359
Db 405 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 425
QY 360 GTCTGGCTATCGCTGATGTGTCTGGGGGTTTATCATATCTCTCTCTCTCTCTCTCTGC 419
Db 425 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyHisIleProLeuHisProAlaAl 445
QY 420 TATGCTCTCATCTTCTTGTGTTCTCTTGACTACCAAGGTATGTGTCGCTTGTCTCTCTC 479
Db 445 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyValAlaArgLeuSerSe 465
QY 480 TACTTCCAGGAACATCAACACGAGCGGGGCGCATGCAAGACCTGCACGACTCTCTCTC 539
Db 465 rAsnSerArgIleIleAsn***GlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe 485
QY 540 AAGGAAACTCTACGTTTCCCTCTTGTGCTGTGTAACAAACCTTCGGACGGAACCTGCACTT 599
Db 485 rArgAsnLeuTyValSerLeuLeuLeuTyLysThrPheGlyArgLysLeuHisLe 505
```

QY 600 GTATTCCCATCCATCATCTGGGCTTTCGCAAGATTCCTATGGGAGTGGGCTCAGTCC 659
 Db 505 ufySerHisProIleLeuGlyPheArglysiIleProMetGlyValGlyLeuSerPr 525
 QY 660 GTTCTCCTGGCTCAGTTACTAGTGCATTTGTTTCAGTGGTTCGTAGGGCTTCCCA 719
 Db 525 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 545
 QY 720 CTGTTTGGCTTTCAGTTATATGATGATGGTATTTGGGGCGAAGTCTGTACAACATCT 779
 Db 545 sCysLeuAlaPheSerTyMetAspAspValValLeuGlyAlaIleValGlnHisLe 565
 QY 780 TGAGTCCCTTTTACTCTATATACCAATTTCTTTTGTCTTTCGGGTATACATTTAAACCC 839
 Db 565 uGluSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 585
 QY 840 TAATAAAACCAACGTTGGGCTACTCCCTTAACTTTCATGGGATATGTAATTGGAAGTTG 899
 Db 585 oAsnLysThrLysArgTyrPglyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 605
 QY 900 GGGTACTTTTACCGCAGGAACATATTGTACTTAAACTCAAGCAATGTTTTCGAAACTGCC 959
 Db 605 pGlyThrLeuProGlnGluHisIleValGlnLysIleLysGlnCysPheArgLysLeuPr 625
 QY 960 TGTAAATAGACCTATTGATTGGAAGATATGTCAAGAAATGTCGGTCTTTGGGCTTTC 1019
 Db 625 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuGlyPheAl 645
 QY 1020 TGCCCTTTTACACATATGGCTATCTGCTGATGATGCTTTATATGATGATATCAATC 1079
 Db 645 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyAlaCysIleGlnSe 665
 QY 1080 TAAGCAGGCTTTCACCTTCTCGCAACTTCAAGGCTTTCGTGTAAACAATATCTGAA 1139
 Db 665 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 685
 QY 1140 CTTTACCCCGTTGCGGCAACGGTCCGCTCTCGCAAGTGTGTTGCTGACGCAACCC 1199
 Db 685 nLeuTyProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 705
 QY 1200 CACTGATGGGCTTCGCATAGCCATCAGCCATGCTGGGCTTTCGCTGCTCTCTCT 1259
 Db 705 oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValAlaProLe 725
 QY 1260 GCGCATCCATCTCGGAACTCTAGCAGCTTGTGTTGCTCGCAGCGGTCTCGAGCAAA 1319
 Db 725 uProIleHisThrAlaGluLeuLeuAlaLysPheAlaArgSerArgSerGlyAlaLys 745
 QY 1320 ACTTATCGGAACCGAACACTCTGTGTCCTCTCTCGGAATATACACTCTCTTCATGGCT 1379
 Db 745 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 765
 QY 1380 GCTAGGTGTGCTGCCAAGTGTCTGCGGGAGCTCTCTGCTGCTGCTGCTGCTGCTGCT 1439
 Db 765 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 785
 QY 1440 GCTGAATCCCGGACGACCGCTCTCGGGCGCGTTCCTACCGTCTCTCTCTCTCTCTCA 1499
 Db 785 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyArgProLeuLeuArg 805
 QY 1500 TCTGCCGTTCGGCCGACCGGCGGACCTCTCTTTTACCGGTCTCTCCCGTATGTGCC 1559
 Db 805 gLeuProPheArgProThrGlyArgThrSerLeuTyAlaValSerProSerValPr 825
 QY 1560 TTCTCATCTCGGACCGTGTGCACTTTCGCTTCTCCTCTGCTGCTGCTGCTGCTGCTGCT 1619
 Db 825 cSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 845
 QY 1620 G 1620
 Db 845 o 845

RESULT 11
 US-08-591-502B-51
 ; Sequence 51, Application US/08591502B
 ; Patent No. 6607727
 ; GENERAL INFORMATION:
 ; APPLICANT: Chisari, Francis V.
 ; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
 ; Lymphocyte Responses to Hepatitis B Virus
 ; NUMBER OF SEQUENCES: 99
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend and Crew LLP
 ; STREET: Two Embarcadero Center, Eighth Floor
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94111-3834
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/591,502B
 ; FILING DATE: 20-May-1996
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/749,540
 ; FILING DATE: 26-AUG-1991
 ; APPLICATION NUMBER: US 07/935,898
 ; FILING DATE: 26-AUG-1992
 ; APPLICATION NUMBER: US 08/100,870
 ; FILING DATE: 02-AUG-1993
 ; APPLICATION NUMBER: WO PCT/US94/08685
 ; FILING DATE: 01-AUG-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Weber, Ellen Lauver
 ; REGISTRATION NUMBER: 32,762
 ; REFERENCE/DOCKET NUMBER: 014740-000230US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 576-0200
 ; TELEFAX: (415) 576-0300
 ; INFORMATION FOR SEQ ID NO: 51:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 842 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: <Unknown>
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 51:
 ; US-08-591-502B-51

Alignment Scores:
 Pred. No.: 2,19e-241 Length: 842
 Score: 2734.50 Matches: 517
 Percent Similarity: 96.86% Conservative: 7
 Best Local Similarity: 95.56% Mismatches: 15
 Query Match: 46.29% Indels: 3
 DB: 2 Gaps: 1

US-10-761-006A-1 (1-3215) x US-08-591-502B-51 (1-842)

QY 1 CTCACCAACATTCACCAAGCTCTGTAGATCCAGGGTGGGGCTATATTTCTCTGC 60
 Db 304 LeuHisHisIleSerProSerProAlaArgSerGlnSerGlyProIlePheSerSer 323
 QY 61 TGTGTGCTCCAGTTCGCGACAGTAAACCTGTTCCGACTTACTGCTCTCCCATATCGTC 120
 Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 343
 QY 121 AATCTTCTCGAGGACTGGGGACCCCTGCACCGAATCGAGAACACCAACATCAGGATTCCT 180
 Db 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363

QY 181 AGGACCCCTGCTCGTTACAGCGGGGTTTTCTCGTTGACAGAAATCCTCAATACC 240
Db |||||
364 ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 383
QY 241 GCAGAGTCTAGACTCTG-GTGAGCTTCTCAATTTCTAGGGGAGCACCCAGTGTC 299
Db |||||
384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
QY 300 CTGGCCAAATTCGAGTCCCACTCCAACTCACTCAACCACTCTTGTCTCCTCAATTT 359
Db |||||
403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuSerSerAsnLe 423
QY 360 GTCCCTGGCTATCGCTGGATGTCTCGCGGGTTTTATCATATCTCTTCATCTGCTGC 419
Db |||||
423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
QY 420 TATGCTCATCTCTTGTGTGTTCTCTCGACTACCAAGGTATGTTCGCCGTTGTCTCTC 479
Db |||||
443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaAlaArgLeuSerSe 463
QY 480 TACTTCCAGGAACATCAACACCAACAGCACGGGCCCATGCAAGACCTGCACGACTCTCTGCTC 539
Db |||||
463 rThrSerArgAsnIleAsnHisGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
QY 540 AAGGAACTCTACGTTTCCCTCTTGTCTGTGTACAAAACCTTCGAGCGGAACCTGCACATT 599
Db |||||
483 rArgAsnLeuTyrValSerLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 503
QY 600 GTATTCCCATCCATCATCTGGGCTTCGCAAGATTCCTATGGGAGTGGCCCTCAGTCC 659
Db |||||
503 uTyrSerHisProIleIleLeuGlyPheArgLysIlePrometGlyGlyGlyLeuSerPr 523
QY 660 GTTCTCTCGCTCAGTTTACTAGTGCATTTGTTTCAGTGGTTCGTAGGGCTTTCCCCCA 719
Db |||||
523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValAlaArgAlaPheProHi 543
QY 720 CTGTTTGGCTTTCAGTTATATGATGATGTGTATGTTGGGGCGAAGTCTGTACAACTCT 779
Db |||||
543 sCysLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLe 563
QY 780 TGAGTCCCTTTTACTCTATTACCAATTTCTTTTGTCTTGGGTATACATTTAAACCC 839
Db |||||
563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
QY 840 TAATAAAACCAACGTTCGGGCTACTCCCTTAACTTCAATGGGATATGTAATTTGGAAGTTG 899
Db |||||
583 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
QY 900 GGGTACTTTACGGAGGACATATTTGTACTAAACTCAAGCATGTTTTCGAAACTGCC 959
Db |||||
603 pGlyThrLeuProGlnGluHisIleValLeuLysIleLysGlnCysPheArgLysLeuPr 623
QY 960 TGTAATAGACTATTGATTTGAAAGTATGCAAGAAATCTGGGTCTTTTGGGCTTTCG 1019
Db |||||
623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
QY 1020 TGCCCTTTTACAAATGTGGCTATCTCGCTTCATGCTTTTATATGATGATATACATC 1079
Db |||||
643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
QY 1080 TAAGCAGGCTTTCATTTCTCGCCAACTTACAGGCCCTTCTGTGTAAACAATATCTGAA 1139
Db |||||
663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuHi 683
QY 1140 CTTTACCCTGGTTCGGCGCAACGTCGGCTCTCTGCCAAGTGTTCGTGACGCAACCCC 1199
Db |||||
683 sLeuTyrProValAlaArg---ArgThrAlaLeuCysGlnValPheAlaAspAlaThrPr 702
QY 1200 CACTGGATGGGCTTGGCCATAGGCCATCAGCGCATGGCTGGAACCTTTCTGGCTCTCTCT 1259
Db |||||
702 oThrGlyTrpGlyLeuAlaIleGlyHisArgMetArgGlyThrPheValAlaProLe 722
QY 1260 GCCGATCCATCTAGCGGAACCTCTAGCAGCTTGTCTGCTGCAGCCGCTTCGGAGCAA 1319

Db |||||
722 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys 742
QY 1320 ACTTATCGGAACCGACAACTCTGTTGCTCTCTCGGAAATACACCTCCTTTCCATGCT 1379
Db |||||
742 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 762
QY 1380 GCTAGGCTGTCTGCCAACTGGATCTCGCGGGACGCTCTTGTCTACGTCCTCGTCGCC 1439
Db |||||
762 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrTyrPheValTyrValProSerAl 782
QY 1440 GCTGAATCCCGGACGACCGCTCTCGGGCGCGTTGGGGCTCTACCGTCCCTTCTTCA 1499
Db |||||
782 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuIleArgProLeuLeuHi 802
QY 1500 TCTCGCTTCGGCGGACGACCGGGCGCACCTCTCTTTACGCGGTCTCCCGTATGTGCC 1559
Db |||||
802 sLeuArgPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 822
QY 1560 TTCTCATCTCGCGGACCGTGTGCATCTTCGCTTTCACCTCTGCACGTCGCATCGAGACACC 1619
Db |||||
822 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaAlaTrpArgProPr 842
QY 1620 G 1620
Db |||||
842 o 842
RESULT 12
US-08-591-502B-50
; Sequence 50, Application US/08591502B
; Patent No. 6607727
; GENERAL INFORMATION:
; APPLICANT: Chisari, Francis V.
; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
; Lymphocyte Responses to Hepatitis B Virus
; NUMBER OF SEQUENCES: 99
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,502B
; FILING DATE: 20-May-1996
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 08/100,870
; FILING DATE: 02-AUG-1993
; APPLICATION NUMBER: WO PCT/US94/08685
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 014740-000230US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 842 amino acids
; TYPE: amino acid
; STRANDBDNES: <Unknown>

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatenIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,502B
FILING DATE: 20-May-1996
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,540
FILING DATE: 26-AUG-1991
APPLICATION NUMBER: US 07/935,898
FILING DATE: 26-AUG-1992
APPLICATION NUMBER: US 08/100,870
FILING DATE: 02-AUG-1993
APPLICATION NUMBER: WO PCT/US94/08685
FILING DATE: 01-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 014740-000230US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 832 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 61:
US-08-591-502B-61

Alignment Scores:

Pred. No.: 7,84e-234 Length: 832
Score: 2652.00 Matches: 494
Percent Similarity: 95.38% Conservative: 22
Best Local Similarity: 91.31% Mismatches: 24
Query Match: 44.90% Indels: 2
Gaps: 0

US-10-761-006A-1 (1-3215) x US-08-591-502B-61 (1-832)

```
QY 1 CTCCACACATTCACCAAGCTCTCTAGATCCAGGCTGAGGCGCTATATTTCCTGTC 60
Db 293 LeuHisAsnLeuProAsnSerAlaArgSerGlnSerGluArgProValPheProCys 312
QY 61 TGGTGGCTCCAGTCCCGAACAGTAAACCCCTGTTCCGACTACTGCTCTCCCATATCGTC 120
Db 313 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerHisIleVal 332
QY 121 AATCTTCTCGAGGACTGGGGACCCCTGCAACCAATGGAGAACACCAATCAGGATTCTCT 180
Db 333 AsnLeuLeuGluAspTrpGlyProCysAlaGluHisGlyGluHisIleArgIlePro 352
QY 181 AGGACCCCTGCTCGTTGACAGCGGGGTTTTCGTGTGACAGAAATCCTCACAATACC 240
Db 353 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 372
QY 241 GCAGAGTCTAGACTCTG-GTGACTTCTCAATTTCTAGGGGGAGCACCCACGTGTC 299
Db 373 AlaGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnTrpArgVal-Se 392
QY 300 CTGGCCAAAATTCGAGTCCCCCACTCCAATCACTCAACCACTCTTGTCTCTCAATTT 359
Db 392 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 412
QY 360 GTCCTGGCTATCGTGGATGTCTGCGCGGTTTATCATATTCCTTCATCTCTGCTGTC 419
Db 412 userTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaAl 432
QY 420 TATGCTCATCTTCTGTGTGTTCTTCTGGACTACCAAGGTATGTGCGCGTTGTGCTC 479
Db 432 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaArgLeuSerSe 452
```

```
QY 480 TACTTCCAGGAACATCAACCAACGAGCGGGCCATGCAAGACCTGCGACGACTCTCTGCTC 539
Db 452 rAsnSerArgIlePheAsnTyrGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe 472
QY 540 AAGGAACTCTACGTTTCCCTCTTGTGCTGTACAAAACCTTCGAGCGGAACCTGACTT 599
Db 472 rArgAsnLeuTyrValSerLeuLeuLeuTyrGlnThrPheGlyArgLysLeuHisLe 492
QY 600 GTATTCCCATCCCATCTCTGGGCTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCC 659
Db 492 uTyrSerHisProIleLeuGlyPheArgLysIlePrometGlyValGlyLeuSerPr 512
QY 660 GTTTCTCTGCTCAGTTTACTAGTGCATTTGTTTCACTGGTTCGTAGGGCTTTCCCA 719
Db 512 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 532
QY 720 CTGTTTGGCTTTCAGTTATATGATGATGTGGTATTTGGGGCGGAGTCTGTACACATCT 779
Db 532 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 552
QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTTCTTTTGTCTTTGGGTATACATTTAAACC 839
Db 552 uGluSerLeuPheThrAlaValThrAsnPhelLeuLeuSerLeuGlyIleHisLeuAsnPr 572
QY 840 TAATAAAACCAACGTTGGGCTACTCCCTTAACTTTCATGGGATATGTAATTGGAAGTTG 899
Db 572 oAsnLysThrLysArgTrpGlyTyrSerLeuHisPheMetGlyTyrValIleGlyCysTy 592
QY 900 GGTACTTTACCGGAGGAACATATGTACTAAACTCAAGCAATGTTTTCGAAAACCTGCC 959
Db 592 rGlySerLeuProGlnAspHisIleIleGlnLysIleLysGluCysPheArgLysLeuPr 612
QY 960 TGTAATAGACCTATTGATTCGAAAGTATGCAAGAAATTCGTGGTCTTTTGGGCTTTCG 1019
Db 612 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 632
QY 1020 TGCCCTTTTACACAATGTGGCTATCTCTGCTTATGCTTATATGATGATATACATC 1079
Db 632 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 652
QY 1080 TAACGAGCTTTCACTTTCTGCCAACCTTCAAGGCTTTCTGTGTAAACAATATCTGAA 1139
Db 652 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 672
QY 1140 CTTTACCCCTTCCCGGCAACGTCGGGTCTCTGCCAAGTGTTCGTGACGCAACCC 1199
Db 672 mLeuTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 692
QY 1200 CACTGGATGGGCTTGGCCATAGGCCATCAGCGCATGGCTGGAACCTTTCTGGCTCTCT 1259
Db 692 oThrGlyTrpGlyLeuValMetGlyHisGlnArgMetArgGlyThrPheLeuAlaArgLe 712
QY 1260 GCCATCCATCTCGGGAACCTCTAGCAGCTTGTGCTCGCAGCCGCTGTGAGCAAA 1319
Db 712 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaAs 732
QY 1320 ACTTATCGGAACCCACCACTCTGTGTCTCTCGGAATACACCTCTTCCATCGCT 1379
Db 732 nIleLeuGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerTyrProTrpLe 752
QY 1380 GCTAGGTGTGCTGCCAATCTGATCCTGCGGGGACGTCCTTTGTCTACGTCCTCGCT 1439
Db 752 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 772
QY 1440 GCTGAATCCCGGAGCACCCGCTCTCGGGCCGTTTGGGGCTCTACCGTCCCTTCTTCA 1499
Db 772 aLeuAsnProAlaAspProSerArgGlyArgLeuGlyLeuSerArgProLeuLeuArg 792
QY 1500 TCTCCGCTTCGGGCGGACCAACCGGCGCTCTCTTTTACGGGTCTCTCCCGTATGCTC 1559
Db 792 gLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaAspSerProSerValPr 812
```


QY 1560 TTCTCATCTGCGGACCGTGTGCACCTTCGCTTACCTCTGCACGTGCGCATGGAGACCACC 1619
DB 812 OSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaIatPArgProPr 832

QY 1620 G 1620
DB 832 0 832

RESULT 14

US-08-591-502B-62

; Sequence 62, Application US/08591502B

; Patent No. 6607727

GENERAL INFORMATION:

; APPLICANT: Chisari, Francis V.

; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T

; Lymphocyte Responses to Hepatitis B Virus

; NUMBER OF SEQUENCES: 99

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP

; STREET: Two Embarcadero Center, Eighth Floor

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA: US/08/591,502B

; FILING DATE: 20-May-1996

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/749,540

; FILING DATE: 26-AUG-1991

; APPLICATION NUMBER: US 07/935,898

; FILING DATE: 26-AUG-1992

; APPLICATION NUMBER: US 08/100,870

; FILING DATE: 02-AUG-1993

; APPLICATION NUMBER: WO PCT/US94/08685

; FILING DATE: 01-AUG-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Weber, Ellen Lauver

; REGISTRATION NUMBER: 32,762

; REFERENCE/DOCKET NUMBER: 014740-000230US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 576-0200

; TELEFAX: (415) 576-0300

; INFORMATION FOR SEQ ID NO: 62:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 832 amino acids

; TYPE: amino acid

; STRANDEDNESS: <Unknown>

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 62:

US-08-591-502B-62

Alignment Scores:

Pred. No.:	2,78e-233	Length:	832
Score:	2646.00	Matches:	495
Percent Similarity:	95.01%	Conservative:	19
Best Local Similarity:	91.50%	Mismatches:	26
Query Match:	44.79%	Indels:	2
DB:	2	Gaps:	0

US-10-761-006A-1 (1-3215) x US-08-591-502B-62 (1-832)

QY 1 CTCACAACTTCCACAGCTGCTAGATCCAGGTGAGGGGCTATATTTCTCTGC 60
DB 293 LeuHisAsnLeuProAsnSerAlaArgSerGlnSerGluArgProValPheProCys 312

QY 61 TGTGGCTCCAGTTCGGAACAGTAAACCCCTGTTCCGACTACTGCTCTCCCATATGTC 120
DB 313 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerLeuIleVal 332
QY 121 AATCTTCTCGAGGACTCGGGACCCCTGCACCCGAAACATGGAGAACACACATCAGATTCT 180
DB 333 AsnLeuLeuGluAspTrpGlyProCysAlaGluHisGlyGluHisIleArgIlePro 352
QY 181 AGGACCCCTGCTCTGTGTACAGCGGGGTTTCTCGTTGACAAGATCTCCACATACC 240
DB 353 ArgThrProSerArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 372
QY 241 GCAGAGTCTAGACTCTG-GTGGACTTCTCTCAATTTCTAGGGGAGACCCACCTGTTC 299
DB 373 AlaGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnTyrArgVal-Se 392
QY 300 CTGGCCAAAATTCGACAGTCCCAACCTCCAATCACTCACCAACCTCTTGTCTCTCAATT 359
DB 392 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuSerSerAsnLe 412
QY 360 GTCTGGGTATCGCTGGATGTCTGGCGGCTTTTATCATATTCCTTCTCATCTGCTGC 419
DB 412 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaAl 432
QY 420 TATGCTCATCTTCTGTTGGTCTCTCTGGACTACCAAGTATGTTGCCGTTTCTCTC 479
DB 432 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaAlaArgLeuSerSe 452
QY 480 TACTTCCAGGAACATCAACACACGACGGGGCCATGCAAGACCTTCGACGACTCTCTGCTC 539
DB 452 rAsnSerArgIleLeuAsnHisGlnHisGlyThrMetProAsnLeuHisAspSerCysSe 472
QY 540 AAGGAACTCTACGTTTCCCTCTGTTGCTGTACAAACCTTCGACGAGAACTCAGCTT 599
DB 472 rArgAsnLeuTyrValSerLeuLeuLeuTyrGlnThrPheGlyArgLysLeuHisLe 492
QY 600 GTATTCCCATCCATCATCTCGGGCTTCGCAAGATTCCTATGGGAGTGGGCTCAGTCC 659
DB 492 uTyrSerHisProIleLeuGlyPheArgLysIlePrometGlyValGlyLeuSerPr 512
QY 660 GTTCTCTGCTCAGTTTACTAGTGCATTTGTTTCTGTTGCTTCTGAGGCTTCTCCCA 719
DB 512 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 532
QY 720 CTGTTTGGCTTTCAGTTATATGATGATGTGGTATGTTGGGGGAGAGTCTGTACACATCT 779
DB 532 sCysLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLe 552
QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTGCTTTTGGGTATACATTTAAACCC 839
DB 552 uGluSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 572
QY 840 TAAATAAACCAACCGTTCGGGCTACTCCCTTAACTTCATGGGATATGTAATTTGAAGTTG 899
DB 572 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlyCysTy 592
QY 900 GGGTACTTTACCGCAGGAAACATATTTGTAATAAATCAAGCAATGTTTTCGAAAACCTGCC 959
DB 592 rGlySerLeuProGlnGluHisIleIleGlnLysIleLysGluCysPheArgLysLeuPr 612
QY 960 TGTAAATAGACCTATTGATTGGAAAGTATGTCAAGAAATGTTGGCTCTTTGGGCTTTGCG 1019
DB 612 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 632
QY 1020 TGCCCCCTTTTACACAATGTGGCTATCTCGCTTATGTCCTTTATATGATGTATGATATCAATC 1079
DB 632 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 652
QY 1080 TAAGCAGGCTTTTACCTTTCTGCCCACTTACAAAGCCTTTTGTGTAAACAATATCTGAA 1139
DB 652 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 672
QY 1140 CCTTTACCCCGTTGCGCGCAACGGTCCGGTCTCTGCCAAGTGTGTTGCTGACGCAACCCC 1199

```
Db      672 nLeuTyzProValAlaGlyGlnArgProGlyLeuCyseGlnValPheAlaAspAlaThrPr 692
Qy      1200 CACTGATGGGCTTGGCCATAGCAGCCATCAGCGATCGCTGGACCTTCTGGCTCTCT 1259
Db      692 oThrGlyTrpGlyLeuAlaMetGlyHisGlnArgMetArgGlyThrPheSerAlaProLe 712
Qy      1260 GCCGATCCATCTCGGAACTCCCTAGCAGCTGTGTTTGGCTCGCAGCGGCTCGAGCAAA 1319
Db      712 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaAs 732
Qy      1320 ACTTATCGGACCGACCAACTCTGTGTCTCTCTCGGAAATACACCTCTCTTCATGGCT 1379
Db      732 nIleIleGlyThrAspAsnSerValValLeuSerArgLysTyThrSerPheProTrpLe 752
Qy      1380 GCTAGGTGTGCTCCCACTGATCGCTCGCGGAGCGCTCTTGTCTACGTCGCGTGGCC 1439
Db      752 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyValProSerAl 772
Qy      1440 GCTGAATCCGCGGACGACCCGCTCTCGGGGCGCTTTCGGGCTCTACCGCTCCCTTCTTCA 1499
Db      772 aLeuAsnProAlaAspAspProSerArgLysGlyArgLeuGlyLeuSerArgProLeuLeuAr 792
Qy      1500 TCTCCGCTTCCGCGCGACGCGGCGGACCTCTCTTTACGCGCTCTCCCGTATGCGC 1559
Db      792 gLeuProPheArgProThrThrGlyArgThrSerLeuTyzAlaAspSerProSerValPr 812
Qy      1560 TTCTCACTCGCGGACCGTGGCACTTCGCTTCACCTCTCGACGTCGCGGAGACACCC 1619
Db      812 oSerHisLeuProAspLeuValHisPheAlaSerProLeuHisValAlaTrpArgProPr 832
Qy      1620 G 1620
Db      832 o 832
```

RESULT 15

```
US-08-591-502B-56
; Sequence 56, Application US/08591502B
; Patent No. 6607727
; GENERAL INFORMATION:
; APPLICANT: Chisari, Francis V.
; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
; Lymphocyte Responses to Hepatitis B Virus
; NUMBER OF SEQUENCES: 99
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,502B
; FILING DATE: 20-May-1996
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 08/100,870
; FILING DATE: 02-AUG-1993
; APPLICATION NUMBER: WO PCT/US94/08685
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Leuwer
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 014740-000230US
```

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 843 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; MOLECULE TYPE: linear
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 56:
US-08-591-502B-56

Alignment Scores:
Pred. No.: 3,46e-233 Length: 843
Score: 2645.00 Matches: 493
Percent Similarity: 95.19% Conservative: 22
Best Local Similarity: 91.13% Mismatches: 25
Query Match: 44.78% Indels: 2
DB: 2 Gaps: 0

US-10-761-006A-1 (1-3215) x US-08-591-502B-56 (1-843)

Qy 1 CTCACAAACATTCACCAAGCTCTGTAGATCCAGGGTGAGGGCCATATATTTCTCTGC 60
Db 304 LeuHisValProProAsnSerSerArgSerGlnSerGlnGlySerValLeuSerCys 323
Qy 61 TGTGTGCTCCAGTCCGGACAGTAAACCTGTTCGACTACTGCTCTCCCATATGTC 120
Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerGluHisCysLeuPheHisIleVal 343
Qy 121 AATCTTCTCGAGGACTGGGACCCCTGCACCGAAACATGGAGAACACCAACATCAGGATTCCT 180
Db 344 AsnLeuIleGluAspTrpGlyProCysAlaGluHisGlyGluHisArgIleArgThrPro 363
Qy 181 AGGACCCCTCTCGTGTTCAGCGGGGTTTTCTCTGTTCACAAAGATCTCTCAATACC 240
Db 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
Qy 241 GCAGACTAGACTCTG-GTGACTTCTCTCAATTTCTAGGGGGAGCACCACGCTTC 299
Db 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnThrArgVal-Se 403
Qy 300 CTGGCCAAAATTCGCGAGTCCCAACCTCCAATCACTCACCAAGCTCTTGTCTTCCAATTT 359
Db 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAspLe 423
Qy 360 GTCTGGCTATCGCTGGATGTGTCTGGCGGTTTTTATCATATTCCTCTTCATCTGCTGC 419
Db 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyHisLeuProLeuHisProAlaAl 443
Qy 420 TATGCTCATCTTCTTGTGTTCTGTGACTACCAAGCTATGTTCGCCGTTGTCTCCTC 479
Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyzrValAlaArgLeuSerSe 463
Qy 480 TACTTCCAGAAACATCAACCAACAGCAGCGGGCCATCAAGACCTCGACCACTCTCTGCTC 539
Db 463 rAsnSerArgIleIleAsnHisGlnHisArgThrMetGlnAsnLeuHisAspSerCysSe 483
Qy 540 AAGGAAACTCTTACGTTTCCCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 599
Db 483 rArgAsnLeuTyzrValSerLeuLeuLeuTyzrLysThrTyzrGlyArgLysLeuHisLe 503
Qy 600 GTATTTCCATCCCATCATCTCTGGGCTTTCAGAGATTCCTATGAGAGTGGGCTCAGTCC 659
Db 503 uTyzrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
Qy 660 GTTTCCTCTCGCTCAGTTTACTAGTGCATTGTTCAGTGTGTGTGTGTGTGTGTGTGTGT 719
Db 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi 543
Qy 720 CTGTTTGGCTTTTCAGTTATATGATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 779
```

Db	543	sCysLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLe	563
QY	780	TGAGTCCCTTTTAACTCTATTACCAATTTCTTTTCTTGGGTATACATTTAAACCC	839
Db	563	uGluSerLeuTyrAlaAlaValThrAsnPheLeuLeuSerLeuGlyLeHisLeuAsnPr	583
QY	840	TAATAAAACCAACGTTGGGGCTACTCCCTTAACCTTCATGGGATATGTAATTCGAAGTTG	899
Db	583	oGlnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr	603
QY	900	GGGTACTTTACCGCAGCAATATGTACTATAAATCTCAACCAATGTTTTCGAAACCTGCC	959
Db	603	pGlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPr	623
QY	960	TGTAATAAGACTTATTGATTGGAAAGATATGTCAAAGAAATTTGGGTCTTTTGGGCTTTGC	1019
Db	623	oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl	643
QY	1020	TGCCCTTTTACAAATGTGGCTATCTGGCTTGATGCTTTATATGATGCATGTATACAATC	1079
Db	643	aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnAl	663
QY	1080	TAAAGACGCTTTCATCTTCTCGCCCACTTACAGGCGCTTTCTGTGTAAACAATATCTGAA	1139
Db	663	AlaGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuThrLysGlnTyrLeuAs	683
QY	1140	CTTTTACCCTTCGCGGCAACGGTCGGGTCTCTGCCAAGTGTTCCTCACGCAACCCC	1199
Db	683	nLeuTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr	703
QY	1200	CACTGGATGGGGCTTGGCCATAGCCATCAGCGCATGGCTGGAACTTTCTGGCTCTCTCT	1259
Db	703	oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValSerProLe	723
QY	1260	GCCGATCCATACTCGGGAACCTCTAGACGCTTGTTCCTCGGACCGGTCTGGAGCAAA	1319
Db	723	uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys	743
QY	1320	ACTTATCGGAACCGACAACCTGTGTGTCTCTCTCGGAAATACACCTTCCTTTCATGGCT	1379
Db	743	sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe	763
QY	1380	GCTAGGGTGTCTGCCAATCGATCTCTCGCGGGACGTCCTTTGTCTACGTCGCGTCCGC	1439
Db	763	uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl	783
QY	1440	GCTGAATCCCGGAGACGACCGCTCTCGGGGGCGTTTGGGGCTCTACCGTCCCTTCTTCA	1499
Db	783	aLeuAsnProAlaAspAspProSerArgLysArgLeuGlyLeuTyrArgProLeuLeuAr	803
QY	1500	TCTGCGTTCGGCGGACCGCGGGCGCACTCTCTTTTACGCGGTCTCCCGTATGTGCC	1559
Db	803	gLeuProTyrArgProThrThrGlyArgThrSerLeuTyrAlaAspSerProSerValPr	823
QY	1560	TTCTCATCTGCGGACCGGTGTCACCTTCGCTTTCACCTCTGCACGTCGCATGGAGACCAC	1619
Db	823	oSerArgLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr	843
QY	1620	G 1620	
Db	843	O 843	

Search completed: December 2, 2005, 00:05:48
Job time : 152.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, (using frame_plus n2p model)

Run on: December 1, 2005, 23:14:00 ; Search time 464.5 Seconds
(without alignments)
6082.251 Million cell updates/sec

Title: US-10-761-006A-1
Perfect score: 5907
Sequence: 1 CTCACACATTCACCAAG.....CCTCAGGCCACGAGTGAA 3215

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-O=/cpn2.1/USPTO.spool.p/US10761006/runat_01122005_113952_26862/app_query.fasta_1.3399
-DB=A_Geneseq -QFW=fastan -SUPFIX=n2p.rag -MINMATCH=0.1 -LOOPEXT=0 -LIST=45
-UNITS=bits -START=1 -END=1 -MATRIX=BLOSUM62 -TRANS=human40.cdi -MODE=LOCAL
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10761006 @CN 1.1 746 @runat_01122005_113952_26862 -NCPU=6 -ICPU=3
-NO WMAP -LARGEOUTRY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -THREDS=6 -DELEXT=7

Database : A_Geneseq 21:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*
9: Geneseq2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2927	49.6	843	3 AAY54044	Aay54044 Amino aci
2	2810	47.6	540	9 ADX40768	Adx40768 HBV polym
3	2810	47.6	540	9 ADX40767	Adx40767 HBV polym
4	2807	47.5	843	9 ADX40772	Adx40772 HBV polym
5	2789	47.2	843	9 ADX40776	Adx40776 HBV polym
6	2778	47.0	845	9 ADX40777	Adx40777 HBV polym
7	2759	46.7	845	2 AAR70065	Aar70065 Hepatitis
8	2750	46.6	843	9 ADX40775	Adx40775 HBV polym
9	2737.5	46.3	539	7 ADM68212	Adm68212 GRP94 pro

10	2734.5	46.3	842	9	ADX40771	ADx40771 HBV polym
11	2688	45.5	825	9	ADX40774	ADx40774 HBV polym
12	2678.5	45.3	827	9	ADX40773	ADx40773 HBV polym
13	2675	45.3	842	4	AAB04708	Aae04708 Hepatitis
14	2675	45.3	842	4	AAG66918	Aag66918 HBV pres1
15	2654	44.9	832	9	ADX40761	ADx40761 HBV polym
16	2651	44.9	832	9	ADX40762	ADx40762 HBV polym
17	2643	44.7	843	9	ADX40760	ADx40760 HBV polym
18	2642	44.7	832	9	ADX40763	ADx40763 HBV polym
19	2628	44.5	832	9	ADX40765	ADx40765 HBV polym
20	2608	44.2	843	3	RAY44348	Aay44348 Human hep
21	2608	44.2	845	9	ADX40758	ADx40758 HBV polym
22	2607	44.1	832	9	ADX40764	ADx40764 HBV polym
23	2607	44.1	845	9	ADX40759	ADx40759 HBV polym
24	2582	43.7	843	9	ADX40766	ADx40766 HBV polym
25	1829	31.0	373	9	ADW42965	Adw42965 ADV (ade)
26	1829	31.0	373	9	ADW42948	Adw42948 ADV (ade)
27	1586	26.8	344	9	AEA08851	Aea08851 Hepatitis
28	1557	26.4	344	9	AEA08850	Aea08850 Hepatitis
29	1552	26.3	303	9	ADX40769	ADx40769 HBV polym
30	1552	26.3	303	9	ADX40770	ADx40770 HBV polym
31	1496	25.3	400	3	AY54045	Aay54045 Amino aci
32	1479	25.0	309	9	AEA08832	Aea08832 Hepatitis
33	1465	24.8	281	1	AAP60163	Aap60163 Subtype a
34	1465	24.8	281	1	AAP60560	Aap60560 Hepatitis
35	1465	24.8	281	1	AAP60617	Aap60617 C-termina
36	1462	24.8	307	9	AEA08833	Aea08833 Hepatitis
37	1461	24.7	281	1	AAP70294	Aap70294 Subtype a
38	1459	24.7	400	2	AAR93801	Aar93801 Hepatitis
39	1459	24.7	400	6	ABR55863	ABr55863 HBV S-pro
40	1459	24.7	400	8	ADN49718	Adn49718 Hepatitis
41	1459	24.7	400	8	ADU74394	Adu74394 HBV S-pro
42	1457	24.7	348	7	ADG76935	Adg76935 Hepatitis
43	1456	24.6	281	2	AAR62870	Aar62870 Hepatitis
44	1456	24.6	281	4	AAM48419	Aam48419 Protein #
45	1452	24.6	389	1	AAP60794	Aap60794 Adr-type

ALIGNMENTS

RESULT 1

RAY54044
ID AAY54044 standard; protein; 843 AA.

XX AC AAY54044;
XX AC
XX 27-MAR-2000 (first entry)
XX DT
XX DE Amino acid sequence of a HBV DNA polymerase protein.
XX KW
XX KW HBV; HBV surface antigen-'S'-133 Oon strain (Met to Thr); DNA polymerase;
XX KW large surface antigen; core protein; transactivating X protein;
XX KW hepatitis vaccine; HBV infection; hepatocellular carcinoma.
XX OS Hepatitis B virus.

XX XX
XX Key Location/Qualifiers
XX FT Misc-difference 194 /note= "encoded by GAG"
XX FT FT
XX FT Misc-difference 195 /note= "encoded by CAA"
XX XX
XX PN WO9966048-A1.
XX XX
XX PD 23-DEC-1999.
XX XX
XX PF 19-JUN-1998; 98WO-SG0000046.
XX XX
XX PR 19-JUN-1998; 98WO-SG0000046.
XX XX
XX PA (GOVE-) GOVERNMENT REPUBLIC SINGAPORE.
XX XX
XX PI Oon CJ, Lim GK, Zhao Y, Chen WN;

ID ADX40768 standard; protein; 540 AA.
 XX AC ADX40768;
 XX DT 21-APR-2005 (first entry)
 XX DE HBV polymerase protein #11.
 XX KW Immune stimulation; polymerase; enzyme.
 XX OS Hepatitis B virus.
 XX PN WO2005012502-A2.
 XX PD 10-FEB-2005.
 XX PF 29-MAR-2004; 2004WO-US009510.
 XX PR 28-MAR-2003; 2003US-0458026P.
 XX PA (EPIM-) EPIMUNE INC.
 XX PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
 XX DR WPI; 2005-132661/14.
 XX
 XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
 PT response comprises identifying variants of a peptide epitope 8-11 amino
 PT acids in length comprising primary anchor residues of the same HLA class
 PT I binding motif.
 XX
 XX Disclosure; Page 380-385; 458pp; English.
 XX
 XX The invention relates to a method of identifying a candidate peptide
 CC epitope which induces an HLA class I CTL response against variants of the
 CC peptide epitope, comprising identifying, from a particular antigen of an
 CC infectious agent, variants of a peptide epitope comprising primary anchor
 CC residues of the same HLA class I binding motif. The method is useful for
 CC identifying a candidate peptide epitope, which induces an HLA class I CTL
 CC response against variants of the peptide epitope. This sequence
 CC represents an HBV polymerase protein used in the scope of the invention.
 XX
 XX Sequence 540 AA;
 SQ

Alignment Scores:
 Pred. No.: 1.68e-214 Length: 540
 Score: 2810.00 Matches: 527
 Percent Similarity: 98.34% Conservative: 5
 Best Local Similarity: 97.41% Mismatches: 8
 Query Match: 47.57% Indels: 2
 DB: 9 Gaps: 0

US-10-761-006A-1 (1-3215) x ADX40768 (1-540)

QY 1 CTCACACATCCACAGCTCTGCTAGATCCAGGGGTGAGGGGCTATATTTCTCTGC 60
 Db 1 LeuHieAsnIleProSerSerAlaArgSerGlnSerGlnSerGlnGlyProIlePheSerCys 20
 QY 61 TGTGTGCTCCAGTTCGGACAGTAACCTGTTCGACTACTCTCTCCCATATCGTC 120
 Db 21 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTrpCysLeuThrHisIleVal 40
 QY 121 AATCTTCTCAGGACTGGGACCCCTGCACCGAACATGGAGAACACCAATCAGGATCTCT 180
 Db 41 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 60
 QY 181 AGACCCCTCTCTGTGTACAGCGGGGTTTTCTCTGTGACAAAGATCTCTCAATACC 240
 Db 61 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 80
 QY 241 GCAGACTAGACTCTG-GTGGACTTCTCTCAATTTCTAGGGGGAGACCCACGCTGTC 299
 Db 81 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 100

QY 300 CTGSCCAAAATTCGAGTCCCACTCCCACTCCCACTCCCACTCCCACTCCCACTCCCACTTT 359
 Db 100 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuSerSerAsnLe 120
 QY 360 GTCTGTGCTATCGTGGATGTGTCTGCGCGCTTTTATCATATTCCTTCTTCCTGCTGC 419
 Db 120 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyHisIleProLeuHisProAlaAl 140
 QY 420 TATGCTCATCTCTTGTGTGTTCTCTGAGCTACCAAGGTATGTGGCGCTTGTCTCTC 479
 Db 140 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyValAlaArgLeuSerSe 160
 QY 480 TACTTCCAGGACATCAACCAACGACGCGGCCCATGCAAGACCTCGACGACTCTCTGCTC 539
 Db 160 rThrSerArgAsnIleAsnTyHisGlyThrMetGlnAsnLeuHisAspSerCysSe 180
 QY 540 AAGGAAACTCTACGTTTCCCTCTCTGCTGTACAAACCTTCGACGCGAAATCTGCACATT 599
 Db 180 rArgAsnLeuTyValSerLeuLeuLeuLeuTyHisThrPheGlyArgGlyLeuHisLe 200
 QY 600 GTATTCCCATCCATCATCTGCGGCTTCGCAAGATTCTTATGGAGTGGGCTCAGTCC 659
 Db 200 uTy-SerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 220
 QY 660 GTTTCCTCTGCTCAGTTTACTAGTGCCTATTTGTTTCTAGTGGTTCGTAGGGCTTCCCCCA 719
 Db 220 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValAlaArgAlaPheProHi 240
 QY 720 CTGTTTGGCTTTCAGTTATATGATGTGTGTATTTGGGGCGAAAGTCTGTACAACATCT 779
 Db 240 sCysLeuAlaPheSerTyMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 260
 QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTTGGGTATACATTTAAACCC 839
 Db 260 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 280
 QY 840 TAAATAAACCAACGTTGGGCTACTCCCTTAACTTCATGGATATGATTTGGAGTTG 899
 Db 280 oAsnLysThrLysArgTrpGlyTy-SerLeuAsnPheMetGlyTyValIleGlySerTr 300
 QY 900 GGGTACTTTTACCGCAGGAAACATATTGTACTAAACCTCAAGCAATGTTTTCGAAAACGTC 959
 Db 300 pGlyThrLeuProGlnGluHisIleValGlnLysLeuLysGlnCysPheArgLysLeuPr 320
 QY 960 TGTAAATAGACCTATTGTATGGAAAGTATGTCAAAGAATTGTGGGTCTTTTGGGCTTTC 1019
 Db 320 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 340
 QY 1020 TGCCCTTTTACAAATGTGGCTATCTCTGCTTGTATGCTTATATGATGTATACATC 1079
 Db 340 aAlaProPheThrGlnCysGlyTy-ProAlaLeuMetProLeuTyAlaCysIleGlnSe 360
 QY 1080 TAAGCAGGCTTTTCACTTTCTCGCCCACTTACAAAGGCTTTCTGTGTAAACAATATCTGAA 1139
 Db 360 rLysGlnAlaPheThrPheSerProThrTyLysAlaPheLeuCysLysGlnTyLysLeuAs 380
 QY 1140 CCTTTACCCCGTTCGCGCAACGGTCCGCTCTCTGCCAAGTGTGTTGTGACCAACCCC 1199
 Db 380 nLeuTyProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 400
 QY 1200 CACTGTAGTGGGCTTGGCCATAGGCCATAGGCCATAGGCCATAGGCCATTTCTGCTCTCT 1259
 Db 400 oThrGlyTyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLe 420
 QY 1260 GCGCATCCATCTCGGGAACCTCTAGCAGCTGTTTGTCTCGCAGCGGTCTCGAGCAAA 1319
 Db 420 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLy 440
 QY 1320 ACTTATCGGAACCGCAACTCTGTGTCTCTCTCGGAATATACACCTCTCTTTCATGGCT 1379
 Db 440 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyThrSerPheProTrpLe 460

```
QY 1380 GCTAGGGTGTGCTGCTCCAACTGGATCTCGCGGGAGCGTCCTTTGTCTACGTCCCGTGGC 1439
D 1460 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrProSerAl 480
QY 1440 GCTGAATCCCGGAGCAGCCGCTCTCGGGGCGGTGGGGCTCTACGTCCTCCCTTCTTCA 1499
D 480 aLeuAsnProAlaAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 500
QY 1500 TCTGCGGTTCCGGCGGACACCGGGCGCACCTCTCTTTACGCGGTCTCCCGTATGTGCC 1559
D 500 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 520
QY 1560 TTCTCATCTCGGACCGGTGCGACTTCGCTTACCTCTGACGTCGTCATGGAGCACACC 1619
D 520 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 540
QY 1620 G 1620
D 540 o 540
RESULT 3
ID ADX40767 standard; protein; 540 AA.
XX ADX40767;
AC ADX40767;
DT 21-APR-2005 (first entry)
XX HBV polymerase protein #10.
DE HBV polymerase; enzyme.
XX Immune stimulation; polymerase; enzyme.
XX Hepatitis B virus.
OS WO2005012502-A2.
PN 10-FEB-2005.
XX 29-MAR-2004; 2004WO-US009510.
XX 28-MAR-2003; 2003US-0458026P.
XX (EPIM-) EPIMMUNE INC.
XX Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX WPI; 2005-132661/14.
XX
PT Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
XX
PS Disclosure; Page 380-385; 458pp; English.
XX
CC The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HBV polymerase protein used in the scope of the invention.
XX
SQ Sequence 540 AA;
Alignment Scores:
Pred. NO.: 1.68e-214 Length: 540
Score: 2810.00 Matches: 527
Percent Similarity: 98.34% Conservative: 5
Best Local Similarity: 97.41% Mismatches: 8
Query Match: 47.57% Indels: 2
DB: 9 Gaps: 0
```

```
US-10-761-006A-1 (1-3215) x ADX40767 (1-540)
QY 1 CTCCACACATTCACCAAGCTCTGCTAGATCCAGGGTGAGGGCCCTATATTTTCTCTGC 60
D 1 LeuHisAsnIleProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCys 20
QY 61 TGGTGGCTCCAGTTCCCGAACAGTAAACCTGTTTCCGACTACTGCTCTCCCATATCTCTC 120
D 21 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 40
QY 121 AATCTTCTCGAGGACTGCGGACCTGCAACCAATGGAGAACACCAATCAGGATTCCT 180
D 41 AsnLeuLeuGluAspTyrGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 60
QY 181 AGGACCCCTGCTGCTGTATACAGCGGGGTTTTCTCGTTGACAAGAATCCTCACAATACC 240
D 61 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 80
QY 241 GCAGAGTCTAGACTCTG- GTGGACTTCTCTCAATTTTCTAGGGGAGCACCACCTGTCT 299
D 81 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 100
QY 300 CTGCCCAAAATTGCGAGTCCCAACCTCCAATCACTCACCAACCTCTTGTCTCTCCAATT 359
D 100 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 120
QY 360 GTCTGGCTATCGCTGGATGCTCTCGGGGTTTTATCATATTTCTTCTTCACTGCTGTC 419
D 120 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 140
QY 420 TATGCTCATCTCTTGTGTTGTTCTCTGGACTACCAAGGTATGTTGCCCTTTGTCTCTC 479
D 140 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 160
QY 480 TACTTCCAGGAACATCAACACACAGCAGCGGGCCATGCAAGACCTGCACGACTCTCTGCTC 539
D 160 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe 180
QY 540 AAGGAAACTCTACGTTTCCCTCTTGTGCTGTACAAAACCTTCGGACGGAACATGCACTT 599
D 180 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 200
QY 600 GTATTTCATCCCATCATCTCTGGGCTTTCGCAAGATTCCTATGGAGTGGGCCCTCAGTCC 659
D 200 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 220
QY 660 GTTCTCTCTGCTCAGTTTACTAGTGCATTTGTTCAGTGGTTCCTAGGGCTTTTCCCCCA 719
D 220 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi 240
QY 720 CTGTTTGGCTTTCAGTTATATGATGATGCTGTTTGGGGCGAAGCTGTGACACATCT 779
D 240 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 260
QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTTCTTTGTCTTTGGGTATACATTTTAAACCC 839
D 260 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 280
QY 840 TAATAAAACCAACCGTTGGGGCTACTCCCTTAACTTCATGGGATATGTAATTGGAAGTTG 899
D 280 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 300
QY 900 GGGTACTTTTACCGCAGGAACATATTGTACTAAACTCAAGCAATGTTTTCGAAACACTGCC 959
D 300 pGlyThrLeuProGlnGluHisIleValGlnLysLeuLysGlnCysPheArgLysLeuPr 320
QY 960 TGTAATATAGACTATTGATTCGAAAGTATGTCAAAGAAATTTGCGGTCTTTTGGGCTTTCG 1019
D 320 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 340
QY 1020 TGCCCCCTTTTACAAATGTGGCTATCTGCTTGTATGCTTATATGCTTATATGATATACAATC 1079
D 1020 TGCCCCCTTTTACAAATGTGGCTATCTGCTTGTATGCTTATATGCTTATATGATATACAATC 1079
```

Db 340 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 360
Qy 1080 TAAGCAGGCTTTACTTTCTGCCCACTTAACAAGGCTTTCTGTGTAACAAATATCTGAA 1139
Db 360 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 380
Qy 1140 CTTTACCGCTTGGCCGCAACGCTCCGCTCTCTGCAAGTGTGCTGACGCAACCC 1199
Db 380 nLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 400
Qy 1200 CACTGATGGGCTGGCCATAGGCATCAGCGCATGCTGGAACCTTCTGGCTCTCT 1259
Db 400 oThrGlyTyrGlyLeuAlaIleGlyHisArgMetArgGlyThrPheValAlaProLe 420
Qy 1260 GCGATCCATCTCGGAACTCTAGCAGCTTTTGTCTGCGCCGCTCTGGAGCAAA 1319
Db 420 uProIleHisThrAlaGluLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys 440
Qy 1320 ACTTATCGAACCGCAACTCTGTGCTCTCTCGGAATACACCTCTTCCATGGCT 1379
Db 440 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProtriple 460
Qy 1380 GCTAGGCTGTGCTCAACTGATCTCTCGCGGACGCTCTTGTCTAGCTCCGCTCGGC 1439
Db 460 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 480
Qy 1440 GCTGAATCCCGGACGACCCCTCTCGGGCCGCTTGGGGCTCTACCGTCCCTCTTCA 1499
Db 480 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 500
Qy 1500 TCTGCGCTTCCGGCCGACACCGGGCGACCTCTCTTACGGGCTCTCCCGTATGTGCC 1559
Db 500 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 520
Qy 1560 TTCTCATCTCGCGGACCGGTGTCACCTTCCTTCACCTGCGACGTCGCATGAGACACC 1619
Db 520 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 540
Qy 1620 G 1620
Db 540 o 540

RESULT 4
ADX40772
ID ADX40772 standard; protein; 843 AA.
AC ADX40772;
XX 21-APR-2005 (first entry)
DT HBV polymerase protein #15.
DE Immune stimulation; polymerase; enzyme.
KW Hepatitis B virus.
OS WO2005012502-A2.
XX 10-FEB-2005.
XX 29-MAR-2004; 2004WO-US009510.
XX 28-MAR-2003; 2003US-0458026P.
XX (EPIN-) EPIMUNE INC.
PA Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX WPI; 2005-132661/14.
XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class

PT I binding motif.
XX Disclosure; Page 380-385; 459pp; English.
XX The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
XX represents an HBV polymerase protein used in the scope of the invention.
SQ Sequence 843 AA;

Alignment Scores:
Pred. No.: 3,31e-214 Length: 843
Score: 2807.00 Matches: 528
Percent Similarity: 98.33% Conservative: 3
Best Local Similarity: 97.78% Mismatches: 8
Query Match: 47.52% Indels: 2
DB: 9 Gaps: 0

US-10-761-006A-1 (1-3215) x ADX40772 (1-843)

Qy 4 CACACATTTCCACCAAGCTCTGCTAGATCCCGAGGTGAGGGCCCTATATTTCTCTGCTGG 63
Db 305 HisAsnIleProProSerSerAlaArgSerGlnSerGlnGlyProIlePheSerCysTrp 324
Qy 64 TGGCTCCAGTTCOGGAACAGTAAACCCCTGTTCCGACTACTGCTCTCTCCCATATCGTCAAT 123
Db 325 TrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleValAsn 344
Qy 124 CTTCTCGAGGACTGGGACCCCTGCAACCAATGAGAGAACACCAATCAGGATTCCTAGG 183
Db 345 LeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIleProArg 364
Qy 184 ACCCTGCTGCTGTACAGGGGGGTTTTCTGCTTGACAGAACTCTCACAATCCCTCACAAATCCGCA 243
Db 365 ThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThrThr 384
Qy 244 GAGTCTAGACTCTG-GTGGACTTCTCTCAATTTCTAGGGGAGACACCCAGTGTTCCTG 302
Db 385 GluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-SerTr 404
Qy 303 GCCAAATTCGCAGTCCCAACCTCCAATCACTCAACCACTCTCTCTCTCCAAATTTGTC 362
Db 404 pProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeuSe 424
Qy 363 CTGGCTATCGCTGGATGTCTGCGGGCTTTTATCATATTCCTCTCTCATCTGCTGCTAT 422
Db 424 rTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAlaMe 444
Qy 423 GCCTCATCTTCTGTTGTTGTTCTCTGGACTACCAAGTATGTTGCCGCTTTGCTCTCTAC 482
Db 444 tProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSerTh 464
Qy 483 TTCAGGAAACATCAACCAACCAAGCCCAATGCAAGACCTGACGACTCTCTGCTCAAG 542
Db 464 rSerArgAsnIleAsnHisGlnHisGlyAlaMetGlnAspLeuHisAspSerCysSerAr 484
Qy 543 GAAACTTACGTTTCCCTCTTGTGCTGTACAAACCTTCGGACCGGAACCTGCACTTGTA 602
Db 484 gAsnLeuTyrValSerLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLeuTyr 504
Qy 603 TTCCCATCCCATCTCTGCTGCTTTTCGCAAGATTCCTATGCGAGTGGGCTCAGTCCGTT 662
Db 504 rSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerProPh 524
Qy 663 TCTCTGCTGCTCAGTTTACTAGTGCCATTTGTTTTCAGTGGTTCGTAGGCTTTCCCCCAGT 722
Db 524 eLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHisCy 544


```
Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 463
Qy 480 TACTTCCAGGAACATCAACACAGCAGCGGCCCATGCAAGACCTGCACGACTCTCTGCTC 539
Db 463 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
Qy 540 AAGGAACCTACAGTTCCCTCTTCTGCTGTACAAAACCTTCGCGAGAACTGCACCTT 599
Db 483 rArgAsnLeuTyrValSerLeuLeuValTyrLysThrPheGlyArgLysLeuHisLe 503
Qy 600 GTATTCCCATCCATCATCTCGGCTTCGCAAGATTCCTATGGAGTGGGCTCAGTCC 659
Db 503 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
Qy 660 GTTCTCTCTGCTCAGTTACTAGTGCCTATTTGTTTCAAGTTCAGTGGTTCAGGGCTTCC 719
Db 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 543
Qy 720 CTGTTTGGCTTTTCACTATATGATGATGTGTATGGGGCGAGTCTGTACACATCT 779
Db 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
Qy 780 TCAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTAAACCC 839
Db 563 uGluSerLeuPheThrSerIleThrAsnPheLeuSerLeuGlyIleHisLeuAsnPr 583
Qy 840 TAATAAAACCAACGTTGGGCTACTCCCTTAACCTTCATGGGATATGTAATTGGAAGTTG 899
Db 583 oAsnLysThrLysArgTyrGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
Qy 900 GGGTACTTTACCGCAGGAACATATTGTACTATAAACTCAAGCAATGTTTCGAAAACCTGCC 959
Db 603 pGlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPr 623
Qy 960 TCTAAATAGACTATTGATTGGAAGTATGCAAGAAATGTTGGGTCTTTTGGCTTTCG 1019
Db 623 oValAsnArgProIleAspIrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
Qy 1020 TGCCCCCTTTTACCAATGCGCTATCTCTGCTTGATGCTTCTTATATGCAATGATACAAATC 1079
Db 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
Qy 1080 TAAGCAGGCTTTCACCTTTCGCAACTTCAAGGCTTCTGTGTAAACAATATCTGAA 1139
Db 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysGlnGlnTyrLeuHi 683
Qy 1140 CTTTATCCCGTTGCGCGCAAGGTCGCTCTCTCCAGTCTTTGCTGAGCAACCC 1199
Db 683 sLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheGlyAspAlaThrPr 703
Qy 1200 CACTGATGGGCTTGGCCATAGCCATCAGCCATGGCTGGAACCTTTCTGCTCTCTCT 1259
Db 703 oThrGlyTyrGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLe 723
Qy 1260 GCGATCCATATCGCGAACTCTAGCAGCTTGTGTTGCTCGCAGCGGCTCGGACAA 1319
Db 723 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys 743
Qy 1320 ACTTATCGGAACCGACACTCTGCTCTCTCTCGGAATACACCTCTCTTCATGGCT 1379
Db 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProThrPle 763
Qy 1380 GCTAGGCTGTGTCGCAACTGAGTCTCTGCGCGGAGCTCTTGTCTAGCTCCGCTCGGC 1439
Db 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
Qy 1440 GCTGAATCCCGCGAGCAGCCGCTCTCGGCGCCGTTTGGGCTCTACCGTCCCTCTTCA 1499
Db 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 803
Qy 1500 TCTGCCGTTCCGCGCAGCAGCGGCGCAGCTCTTTTACGCGGCTCTCCCGTATGCGC 1559
Db 803 sLeuProPheGlnProThrThrGlyArgThrSerLeuTyrAlaValAsnSerProSerValPr 823
```

```
Qy 1560 TTCTCATCTCGCGACCGCTGTGCACTTCGCTTCACTCTGCACTCGCATCGAGACCACC 1619
Db 823 oSerHisLeuProValArgValHisPheAlaSerProLeuHisValAlaIleTrpArgProPr 843
Qy 1620 G 1620
Db 843 o 843
RESULT 6
ID ADX40777 standard; protein; 845 AA.
XX ADX40777;
XX 21-APR-2005 (first entry)
XX HBV polymerase protein #20.
XX Immune stimulation; polymerase; enzyme.
XX Hepatitis B virus.
XX WO2005012502-A2.
XX 10-FEB-2005.
XX 29-MAR-2004; 2004WO-US009510.
XX 28-MAR-2003; 2003US-0458026P.
XX (EPIM-) EPIMUNE INC.
XX Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX WPI; 2005-132661/14.
XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
XX response comprises identifying variants of a peptide epitope 8-11 amino
XX acids in length comprising primary anchor residues of the same HLA class
XX I binding motif.
XX Disclosure; Page 380-385; 458pp; English.
XX The invention relates to a method of identifying a candidate peptide
XX epitope which induces an HLA class I CTL response against variants of the
XX peptide epitope, comprising identifying, from a particular antigen of an
XX infectious agent, variants of a peptide epitope comprising primary anchor
XX residues of the same HLA class I binding motif. The method is useful for
XX identifying a candidate peptide epitope, which induces an HLA class I CTL
XX response against variants of the peptide epitope. This sequence
XX represents an HBV polymerase protein used in the scope of the invention.
XX Sequence 845 AA;
Alignment Scores:
Pred. No.: 6,74e-212 Length: 845
Score: 2778.00 Matches: 521
Percent Similarity: 97.41% Conservative: 6
Best Local Similarity: 96.30% Mismatches: 13
Query Match: 47.03% Indels: 2
DB: Gaps: 0
US-10-761-006A-1 (1-3215) x ADX40777 (1-845)
```

```
Qy 1 CTCCCAACATTCACCAAGCTCTGCTAGATCCAGGTGAGGGCTATATTTCTCTGC 60
Db 306 LeuHisAsnIleProProSerSerAlaArgSerGlnSerGlyPro***PheSerCys 325
Qy 61 TGTGTGCTCCAGTTCCGGAACAGTAAACCTGTTCGACTACTGCTCTCCCATATGCTC 120
Db 326 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerHisIleVal 345
```


CC used, particularly ex vivo, to stimulate CTL cells. These cells can be
 CC reintroduced into patients who have chronic or acute HBV infections or
 CC are carriers, especially in treatments to prevent conversion from acute
 CC to chronic infections. (Updated on 25-MAR-2003 to correct PN field.)

XX
 SQ Sequence 845 AA;

Alignment Scores:

Pred. No.: 2,19e-210 Length: 845
 Score: 2759.00 Matches: 516
 Percent Similarity: 97.23% Conservative: 10
 Best Local Similarity: 95.38% Mismatches: 14
 Query Match: 46.71% Indels: 2
 DB: 2 Gaps: 0

US-10-761-006A-1 (1-3215) x AAR70065 (1-845)

```

QY 1 CTCACAACTTCCACCAAGCTCTGTAGATCCAGGGTGAGGGCGGCTATATTTCTGTC 60
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
306 LeuHisAsnIleProProAsnSerAlaArgSerGlnSerGluGlyProValPheSerCys 325
QY 61 TGTGTGCTCCAGTTCGGAACAGTAACCTCTGTTCGACTACTGCTCTCCCATATCGTC 120
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
326 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 345
QY 121 AATCTTCTCAGAGACTGGGGACCTCGCACGACATGGAGAACACACATCAGGATTCT 180
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
346 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 365
QY 181 AGGACCCCTGCTGTTACAGCGGGGTTTTCTGTTGACAAGAACTCTCACATACC 240
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
366 ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 385
QY 241 GCAGAGCTTAGACTCTG-GTGACTTCTCTCAATTTTCTAGGGGAGACCCACGCTGTT 299
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
386 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrArgVal-Se 405
QY 300 CTGGCCAAATTCGAGTCCCAACCTCCAACTCACACACCTCTGCTCTCCCAATTT 359
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
405 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 425
QY 360 GTCTGCTGCTAGTGTGCTGGGGGTTTTATCATATTCCTTCATCTCTGCTCTGCT 419
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
425 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 445
QY 420 TAGCTCTCATCTTGTGTGTTCTTCTGACTTACCAAGTATGTTGCCCGTTTGTCTTC 479
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
445 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaArgLeuSerSe 465
QY 480 TACTTCCAGAACATCAACCAACGACGAGCGGGCCATGCAAGACTGACGACTCTGCTC 539
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
465 rAsnSerArgIleAlaAsnTyrGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe 485
QY 540 AAGGAACCTAGCTTCTCTCTGTTGCTGTGTACAAAACCTTCGGACGGAACTGCACCT 599
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
485 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 505
QY 600 GTATTCCATCCATCTCTGGGCTTTCGCAAGATTCTTATGGGAGTGGGCTCAGTCC 659
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
505 uTyrSerHisProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 525
QY 660 GTTTCCTCTGGCTCAGTTTACTAGTCCATTTGTTCAGTGGTTCGTAAGGCTTCCCCCA 719
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
525 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHI 545
QY 720 CTGTTTGGCTTTCAGTTATGATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 779
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
545 sCysLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLe 565
QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTCTTCTTCTTCTTCTTCTTCTTCT 839
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
565 uGluSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 585

```

```

QY 840 TAAATAAACCAACGTTGGGCTACTCCCTTAACTTTCATGCGATATGTAATTGGAAGTTG 899
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
595 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 605
QY 900 GGGTACTTTACCGCAGGACATATTGTACTAAACTCAAGCAATTTTTCGAAAACTGCC 959
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
605 pGlyThrIleProGlnGluHisIleValGlnLysIleLysGlnCysPheArgLysLeuPr 625
QY 960 TGTAAATAGACTATTGATTGGAAGATGTCAAAGAATTGGGCTCTTTGGGCTTTCG 1019
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
625 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 645
QY 1020 TGCCCTTTTACACAATGTGGCTATCTCGCTTGTGCTTGTATATGATGATATACAATC 1079
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
645 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 665
QY 1080 TAAGCAGGCTTTTCACTTCTCGCCCAACTTACAAGGCTTCTTGTGTAAACAATATCTGAA 1139
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
665 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 685
QY 1140 CCTTTACCCCGTTGGCCGCAACGCTCGGCTCTCTGCCAAGNTTGTGCTGACGCAACCC 1199
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
685 nLeuTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 705
QY 1200 CACTGATGGGCTTGGCCATAGGCCATCAGCGCATGGCTGGAACCTTTCTGGCTCTCT 1259
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
705 oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValAlaProLe 725
QY 1260 GCGGATCCATCTACGCGAACTCTTAGCAGCTTGTGTTGCTCGCAGCGGCTCTGGAGCAA 1319
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
725 uProIleHisThrAlaGluLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys 745
QY 1320 ACTTATCGGAACCGCAACTCTGTTGCTCTCTCGGAATAATACACTCTTTCATGGCT 1379
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
745 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 765
QY 1380 GCTAGGCTGTCTGCCCAACTGATCTCGCGGAGCTCTTGTCTCTAGCTCCCGTCGGC 1439
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
765 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 785
QY 1440 GCTGAATCCCGCGGAGCAGCCCGCTCTCGGGCGCTTTTGGGGCTCTACCGTCCCTTCTTCA 1499
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
785 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuAr 805
QY 1500 TCTGCGCTTCCGCGGACCAACGCGGCGCAGCTCTCTTTACGCGGCTCTCCCGGTATGTC 1559
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
805 gLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 825
QY 1560 TTCTCATCTGCGGACCGGTGTCACCTTCGCTTCACCTCTGACGCTGCGATGGAGACCACC 1619
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
825 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 845
QY 1620 G 1620
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
845 o 845

```

RESULT 8

ADX40775
 ID ADX40775 standard; protein; 843 AA.
 XX AC ADX40775;
 XX DT 21-APR-2005 (first entry)
 XX DE HBV polymerase protein #18.
 XX KW Immune stimulation; polymerase; enzyme.
 XX OS Hepatitis B virus.
 XX PN WO2005012502-A2.
 XX PD 10-FEB-2005.

XX 29-MAR-2004; 2004WO-US009510.
XX PF
XX 28-MAR-2003; 2003US-0458026P.
XX PR
XX PA (EPIM-) EPIMMUNE INC.
XX
XX PI Baker DM, Livingston BD, Cheenut RW, Sette A, Newman MJ;
XX WPI; 2005-132661/14.
XX
XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
XX response comprises identifying variants of a peptide epitope 8-11 amino
XX acids in length comprising primary anchor residues of the same HLA class
XX I binding motif.
XX
XX Disclosure; Page 380-385; 458pp; English.
XX
XX The invention relates to a method of identifying a candidate peptide
XX epitope which induces an HLA class I CTL response against variants of the
XX peptide epitope, comprising identifying, from a particular antigen of an
XX infectious agent, variants of a peptide epitope comprising primary anchor
XX residues of the same HLA class I binding motif. The method is useful for
XX identifying a candidate peptide epitope, which induces an HLA class I CTL
XX response against variants of the peptide epitope. This sequence
XX represents an HBV polymerase protein used in the scope of the invention.
XX
XX Sequence 843 AA;
XX
XX Alignment Scores:
XX Pred. No.: 1.14e-209 Length: 843
XX Score: 2750.00 Matches: 517
XX Percent Similarity: 97.04% Conservative: 8
XX Best Local Similarity: 95.56% Mismatches: 15
XX Query Match: 46.55% Indels: 2
XX DB: 9 Gaps: 0
XX
XX US-10-761-006A-1 (1-3215) x ADX40775 (1-843)
XX
XX QY 1 CTCACAACTCCACCAAGCTCTGCTAGATCCCGAGGTGAGGGCCCTATATTTCTCTGC 60
XX DB LeuHisAsnIleProSerSerAlaArgSerLysSerGluGlyProLeuPheProCys 323
XX
XX QY 61 TGGTGGCTCCAGTTCGGGAACAGTAACCTGTTCCGACTACTGCTCTCCCATATCGTC 120
XX DB TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 343
XX
XX QY 121 AATCTCTCGAGGACTGGGACCTCGACCGAATCGGAGACACACATCAGATTCCT 180
XX DB AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
XX
XX QY 181 AGGACCCCTGCTGCTGTTACAGGGGGGTTTTCTCGTTGACAGAATCTCCACAAATCC 240
XX DB ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 383
XX
XX QY 241 GCAGAGCTAGACTCG- GTGAGCTTCTCTCAATTTCTAGGGGAGCACCACGTTTC 299
XX DB ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnThrGlnVal-Se 403
XX
XX QY 300 CTGGCCAAAATTCAGTCCCACTCCCAATCACTACCAACCTCTGTGCTCCCAATTT 359
XX DB rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
XX
XX QY 360 GTCTGTGCTATCGCTGATGTGTCTGGGGGTTTTATCATATTCCTCTTCATCTCTGCTGC 419
XX DB uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
XX
XX QY 420 TAGCCCTCARTCTTCTGTTGGTTCTTCTGGAATACCAAGGTATGTCCTCCGTTTGTCTC 479
XX DB aMetProHisLeuLeuValGlySerSerGlyLeuGlnArgTyrValAlaArgLeuSerSe 463
XX
XX QY 480 TACTTCCAGGAACATCAACACGACGCGGGCCATCGAAGACTCGACGACTCTCTGCTC 539
XX
XX

DB 463 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
QY 540 AAGGAAACTCTACGTTTCCCTCTTGTGTCTGTACAAAACCTTCGGACGGAACCTGCACTT 599
DB 483 rLysHisLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisIle 503
QY 600 GTATTCCCATCCCATCATCTCTGGGCTTTCGCAAGATTCCTATGGAGTGGGCCCTCAGTCC 659
DB 503 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
QY 660 GTTCTCTCGCTCAGTTTACTAGTGCATTTGCTCAGTGTGCTTGTGAGGGCTTTCCCCCA 719
DB 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValCysArgAlaPheProHi 543
QY 720 CTGTTTCTGCTTTCAGTTATATGATGTGGTATTGGGGCGAAGTCTGTACCAATCT 779
DB 543 sCysLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisIle 563
QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCC 839
DB 563 uGluSerLeuPheThrSerIleThrAsnPheMetLeuSerLeuGlyIleHisLeuAsnPr 583
QY 840 TAATAAAACCAACGTTGGGGCTACTCCCTTAACCTTTCATGGGATATGTAATGGAAGTTG 899
DB 583 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
QY 900 GGGTACTTTTACCGCAGCAACATATTGTACTAAAACTCAAGCAATGTTTTCGAAACCTGCC 959
DB 603 pGlyThrLeuProGlnGluHisIleValGlnLysLeuLysGlnCysPheArgLysLeuPr 623
QY 960 TGTAAATAGACCTATTGATTTGGAAAGTATGTCAAGAAATGTGGGTCTTTTGGGCTTTCG 1019
DB 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
QY 1020 TGCCCTTTTACAAATGTGGCTATCTGCTTGCATGCTTATATGATGATATACAAATC 1079
DB 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
QY 1080 TAAGCAGGCTTTCACTTTCTCGCCAACTTACAAAGCCTTTCTGTGTAAACAATATCTGAA 1139
DB 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 683
QY 1140 CTTTACCCGCTGTCGGGCAACGGTCCGCTCTCTGCAAGTGTTCGTGACGCAACCC 1199
DB 683 nLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
QY 1200 CACTGGATGGGGCTTGGCCATAGCCCATAGCCGATGCTGGAAACCTTCTGCGCTCTCTCT 1259
DB 703 oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValAlaProLe 723
QY 1260 GCGATCCATACTGCGGAACCTCTTAGCAGCTTGTGTTTGTCTGCGACGGCTCTGGAGCAA 1319
DB 723 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaAth 743
QY 1320 ACTTATCGGAACCGACAACCTCTGTGTCTCTCTCGGAAATACACCTCTTTCATCGGCT 1379
DB 743 rLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 763
QY 1380 GCTAGGTGTGCTGCCAATCGGATCTGCGGGGACGCTCTTGTCTACGTCCCGTGGGC 1439
DB 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
QY 1440 GCTGAATCCCGCGACGACCCGCTCTCGGGCCGCTTGTGGGCTCTTACCGTCCCTCTTCA 1499
DB 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuAr 803
QY 1500 TCTGCCGCTTCCGGCCGACACGCGGGCGCACCTCTCTTTACGCGGTCTCCCGTATGTC 1559
DB 803 gLeuSerPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
QY 1560 TTCTCATCTCGCGGACCGTGTGCATCTCGCTTACCTCTCTGACGTCTCGATCGGACGAC 1619
DB 823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 843

Db	380	sLeuTyProValAlaArg---ArgThrAlaLeuCysGlnValPheAlaAspAlaThrPr	399
Qy	1200	CAC TGGATGGGCGCTGGCCATAGCGCCATAGCGCATGGCTGGAACCTTCTGGGCTCCTCT	1259
Db	399	oThrGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLe	419
Qy	1260	GCCGATCATACTCGGAACCTCTAGCAGCTGTGTTTCTCGCAGCGGCTCTGGAGCAAA	1319
Db	419	uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLy	439
Qy	1320	ACTTATCGGACCGACAACTCTGTGTCTCTCTCGGAATATACACTCCCTTTCATCGCT	1379
Db	439	sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyThrSerPheProTrpLe	459
Qy	1380	GCTAGGCTGTCTGCCAACTCGATCTCGCGGGACGCTCTTCTTCTACGTCGCGTCGGC	1439
Db	459	uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrTyPheValTyValProSerAl	479
Qy	1440	GCTGAATCCCGCGACGACCGCTCTCGGGCGCGTTTGGGCTCTACCGTCCCTCTTCTCA	1499
Db	479	aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuIleArgProLeuLeuHi	499
Qy	1500	TCTGCGCTTCGGCGGACCGGCGGCGACCTCTCTTTACGCGCTCTCCCGGTATGCGC	1559
Db	499	sLeuArgPheArgProThrThrGlyArgThrSerLeuTyAlaValSerProSerValPr	519
Qy	1560	TTCTCATCTGCGGACCGTGTGCATCTCGCTTCACTCTGCACGTCGATGGAGACCA	1619
Db	519	oSerHisLeuProAspArgValHisPheAlaSerProLeuHi sValAlaTrpArgProPr	539
Qy	1620	G 1620	
Db	539	o 539	

RESULT 10

ADX40771
ID ADX40771 standard; protein; 842 AA.

DT 21-APR-2005 (first entry)

HBV polymerase protein #14.

Immune stimulation; polymerase; enzyme.

OS Hepatitis B virus.

PN WO2005012502-A2.

PD 10-FEB-2005.

29-MAR-2004; 2004WO-US009510.

PR 28-MAR-2003; 2003US-0458026P.

PA (EPIM-) EPIMUNE INC.

PI Baker DM, Livingston BD, Chesnut RW, Sette A. Newman MJ:

DR WPI; 2005-132661/14.

Identifying a candidate peptide epitope, which induces a HLA class I CTL response comprises identifying variants of a peptide epitope 8-11 amino acids in length comprising primary anchor residues of the same HLA class I binding motif.

PS Disclosure; Page 380-385: 458pp: English.
XX

The invention relates to a method of identifying a candidate peptide epitope which induces an HLA class I CTL response against variants of the peptide epitope, comprising identifying, from a particular antigen of an

CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HBV polymerase protein used in the scope of the invention.

SQ Sequence 842 AA;

Alignment Scores:		
Prod. No.:	1.959-208	Length:
Score:	2734.50	Matches:
Percent Similarity:	96.88%	Conservative:
Best Local Similarity:	95.55%	Mismatches:
Query Match:	46.29%	Indels:
DB:	9	Gaps:

US-10-761-006A-1 (1-3215) x ADX40771 (1-842)

Qy	1	CTCCAGCAACATTCACCAAGCTCTGCTAGATCCCAAGGTCAGGGGGCTATATTTTCCTGC	60
Db	304	LeuHisHisLeuSerProAlaArgSerGlnSerGluGlyProIlePheSerSer	323
Qy	61	TGTGGCTCCAGTTCGGAACAGATAAACCTGTTCGCACTACTGCCTCTCCCATATCGTC	120
Db	324	TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal	343
Qy	121	AATCTTCTCAGGAGCTGGGACCCTGCACCCGAAATGAGAGAACACAAATCATCAGGATTCCT	180
Db	344	AsnLeuLeuGluAspTTPGlyProCysThrGluHisGlyGluHisAsnIleargIlePro	363
Qy	181	AGNACCCCTGCTGTTTACAGGGCGGGTTTTTCTCGTTTGACAGAAATCCTCACAAATCC	240
Db	364	ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr	383
Qy	241	GCAGAGTCTAGACTCTG-GTGGACTTCTCAATTTTCTAGGGGGAGACCCACCATGTGTC	299
Db	384	ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Ser	403

QY 840 TAATAAAACCAACGTTGGGCTACTCCCTTAACCTTCAATGGGATATGAATGGAGTTG 899
 Db |||||
 QY 593 cAaenLysThrLysAsgTTPGlyTyrSerLeuAenPheMetGlyTyrValIleGlySerTr 603
 Db |||||
 QY 900 GGGTACTTTACCGCAGGACATATGTACTATAAATCAAGCAATGTTTTCGAAACTGCC 959
 Db |||||
 QY 603 pGlyThrLeuProGlnGluHisIleValLeuLysIleLysGlnCysPheArgLysLeuPr 623
 Db |||||
 QY 960 TGTAAATAGACCTATTGATTGGAAGATATGTCAAGAAATTTGGGCTTTTGGGCTTTGC 1019
 Db |||||
 QY 623 oValaenArgProIleAsgTTPlysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
 Db |||||
 QY 1020 TGCCCTTTTACCAATATGGCTATCTCGCTTGTATGCTTGTATATGATGATATCAATC 1079
 Db |||||
 QY 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
 Db |||||
 QY 1080 TAAGCAGGCTTTCACCTTTCTGCCCACTTACAGGCTTTCGTGTAAACATATCTGAA 1139
 Db |||||
 QY 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuHi 683
 Db |||||
 QY 1140 CTTTACCCCGTGGCCGCAACGGTCCGCTCTCTGCCAAGTGTGCTGAGCAACCC 1199
 Db |||||
 QY 683 sLeuTyrProValAlaArg--ArgThrAlaLeuCysGlnValPheAlaAsgPalaThrPr 702
 Db |||||
 QY 1200 CACTGGATGGGCTTGGCCATAGGCCATCAGCGCATGGCTGGAACCTTTCTGGCTCTCT 1259
 Db |||||
 QY 702 oThrGlyTTPGlyLeuAlaIleGlyHisArgMetArgGlyThrPheValAlaProLe 722
 Db |||||
 QY 1260 GCGGATCCATCTAGCGGAACCTCTAGCAGCTTGTGTTGCTCGCAGCGGCTCTGGAGCAA 1319
 Db |||||
 QY 722 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys 742
 Db |||||
 QY 1320 ACTTATCGGAACCGCAACTCTGTTCTCTCTCGGAATACACCTCTTCCATGGCT 1379
 Db |||||
 QY 742 sLeuIleGlyThrAaspAsnSerValValLeuSerArgLysTyrThrSerPheProTrie 762
 Db |||||
 QY 1380 GCTAGGCTGTGTCGCAACTGGAATCTGCGCGGACGTCCTTTGCTACGTCGCGTCGCG 1439
 Db |||||
 QY 762 uLeuGlyCysAlaAlaAasnTrieLeuArgGlyThrTyrPheValTyrValProSerAl 782
 Db |||||
 QY 1440 GCTGAATCCCGGAGCAGCCGCTCTCGGGCCGTTTGGGGCTCTACCGTCCCTCTTCCA 1499
 Db |||||
 QY 782 aLeuAenProAlaAaspProSerArgGlyArgLeuGlyLeuIleArgProLeuLeuHi 802
 Db |||||
 QY 1500 TCTGCGGTTCCGCGCCGACCAACGCGGCACCTCTCTTTAGCGGCTCTCCCGTATGTCG 1559
 Db |||||
 QY 802 sLeuArgPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 822
 Db |||||
 QY 1560 TTCTCATCTCGCGGACCGTGTGCACTTCGCTTCCCTCTGACGTCGCAATGGAGACACC 1619
 Db |||||
 QY 822 oSerHisLeuProAaspArgValHisPheAlaSerProLeuHisValAlaIleArgProPr 842
 Db |||||
 QY 1620 G 1620
 Db 842 o 842
 RESULT 11
 ID ADX40774 standard; protein; 825 AA.
 AC
 AC ADX40774;
 DT 21-APR-2005 (first entry)
 XX HBV polymerase protein #17.
 DE Immune stimulation; polymerase; enzyme.
 KW Hepatitis B virus.
 OS
 XX W02005012502-A2.
 PN
 XX

PD 10-FEB-2005.
 XX 29-MAR-2004; 2004WO-US009510.
 PF
 XX 28-MAR-2003; 2003US-0458026P.
 PR
 XX (EPIIM-) EPIIMUNE INC.
 PA
 XX Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
 PI WPI; 2005-132661/14.
 DR
 XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
 PT response comprises identifying variants of a peptide epitope 8-11 amino
 PT acids in length comprising primary anchor residues of the same HLA class
 PT I binding motif.
 XX
 PS Disclosure; Page 380-385; 458pp; English.
 CC The invention relates to a method of identifying a candidate peptide
 CC epitope which induces an HLA class I CTL response against variants of the
 CC peptide epitope, comprising identifying, from a particular antigen of an
 CC infectious agent, variants of a peptide epitope comprising primary anchor
 CC residues of the same HLA class I binding motif. The method is useful for
 CC identifying a candidate peptide epitope, which induces an HLA class I CTL
 CC response against variants of the peptide epitope. This sequence
 CC represents an HBV polymerase protein used in the scope of the invention.
 XX
 SQ Sequence 825 AA;
 Alignment Scores:
 Pred. No.: 9,766-205 Length: 825
 Score: 2688.00 Matches: 504
 Percent Similarity: 95.92% Conservative: 13
 Best Local Similarity: 93.51% Mismatches: 21
 Query Match: 45.51% Indels: 2
 DB: 9 Gaps: 0
 US-10-761-006A-1 (1-3215) x ADX40774 (1-825)
 QY 7 AACATTCCACCAAGCTCTCTAGATCCAGGGTGAGGGGCGCTATATTTCTCTGCTGGTGG 66
 Db |||||
 QY 288 HisLeuSerThrThrLysArgGlnSerSerSerGlyHisAlaValGluThrCysTrpTrp 307
 Db |||||
 QY 67 CTCGAGTTCGGAAACAGTAAACCTGTTCCGACTACTGCTCTCCCATATCGTCAATCTT 126
 Db |||||
 QY 308 LeuGlnPheArgAasnSerLysProCysSerAspTyrCysLeuThrHisIleValAasnLeu 327
 Db |||||
 QY 127 CTCGAGGACTGGGGCCCTGACCGAATGAGGACACAAACATCAGGATTCCTAGGACC 186
 Db |||||
 QY 328 LeuGluAaspTTPGlyProCysThrGluHisGlyGluHisAasnIleArgIleProArgThr 347
 Db |||||
 QY 187 CTGCTCTGTTTACAGCGGGGTTTCTCTGTTGACAAGAATCTCTCAATATCCGACAG 246
 Db |||||
 QY 348 ProAlaArgValThrGlyGlyValPheLeuValAspLysAasnProHisAasnThrGlu 367
 Db |||||
 QY 247 TCTAGACTCTG-GTGACTTCTCTCAATTTCTTAGGGGGAGACCCACGTTGTTCTGGCC 305
 Db |||||
 QY 368 Ser***LeuValValAaspPheSerGlnPheSerArgGlySerThrGlnVal-SerTrpPr 387
 Db |||||
 QY 306 AAAATTGCGAGTCCCAACCTCAATCCTCAACCACTCTCTCTCTCTCTCTCTCTCTCT 365
 Db |||||
 QY 387 oLysPheAlaValProAasnLeuGlnSerLeuThrAasnLeuLeuSerAasnLeuSerTr 407
 Db |||||
 QY 366 GCTATCTGCTGGATGTTCTGCGCGGTTTATCATATTTCTCTCTCTCTCTCTCTCTCT 425
 Db |||||
 QY 407 pLeuSerLeuAaspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAlaMetPr 427
 Db |||||
 QY 426 TCATCTCTGTTGGTTCCTCTCGACTACCAAGGTATGTTGCCGTTGCTCTCTCTCTCT 485
 Db |||||
 QY 427 oHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSerThrSe 447
 Db |||||
 QY 486 CAGGAACATCAACCAACGACCGGGGCCATGCAAGACCTGCACGACTCTCTCTCAAGAA 545

Db 447 rArgAsnIleAsn***GlnHisGlyThrMetGlnAspLeuHisAspSerCysSerLysHis 467
Qy 546 ACTTACGTTTCCCTCTTGTGCTGATCAAAACCTTCGAGCGGAACATGCACTTGTATTTC 605
Db 467 sLeuTyRValSerLeuLeuLeuTyRLeuThrPheGlyArgLysLeuHisLeuTyRSe 487
Qy 606 CCATCCCATCATCTCTGGGCTTCGCAAGATTCCTATGGAGTGGGCCCTCAGTCCGTTTCT 665
Db 487 rHisProIle***LeuGlyPheArgLysIleProMetGlyValGlyLeuSerProPheLe 507
Qy 666 CCTGCTCAGTTTACTAGTGCATTTGTTAGTGGTTCGTAGGCTTCCGCCACTGTTT 725
Db 507 uLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHisCysLe 527
Qy 726 GGCCTTTCAGTTATATGATGATGTGTTATGGGGCGGAAGTCTGTACAAATCTTGTAGTC 785
Db 527 uAlaPheSerTyRMetAspAspValValLeuGlyAlaLysSerValGlnHisLeuGluAl 547
Qy 786 CTTTTTTACCTCTATTACCAATTTCTTTTCTTTGGGTATATACATTTAAACCTTAATAA 845
Db 547 aLeuTyRThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnProAsnLys 567
Qy 846 AACCAACGTTGGGCTACTCCCTTACTTCACTTATGGGATATGTAATGGAAGTTGGGTAC 905
Db 567 sThrLysArgTrpGlyTyRSerLeuAsnPheMetGlyTyRValIleGlySerTrpGlyTh 587
Qy 906 TTTTACCGCAGGAACATATTGTACTAAACTCAAGCAATGTTTTCGAAACTGCCTGTAAA 965
Db 587 rLeuProGlnAspHisIleValGlnLysLeuLysGlnCysPheArgLysLeuProValAs 607
Qy 966 TAGACCTATTGATGGAAGATGTCAAAGATTTGTGGGTCTTTTGGGCTTGTGCGCCC 1025
Db 607 nArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAlaAlaPr 627
Qy 1026 TTTTACCAAGTGGCTATCTGCTTGCCTTGCATGCTTATATCATGATATACATCTAAGCA 1085
Db 627 oPheThrGlnCysGlyTyRProAlaLeuMetProLeuTyRAlaCysIleGlnSerLysGl 647
Qy 1086 GCGTTTCACTTTTCGCCAACTTCAAGCGCTTCTGTGTAAACAATATCTGAACCTTTA 1145
Db 647 nAlaPheThrPheSerProThrTyRLeuAlaPheLeuCysLysGlnTyRLeuAsnLeuTy 667
Qy 1146 CCCGTTGCCCGCAACCGTCCGGTCTCTGCAAGTGTTCGCAAGTGTTCGCAACCCCACTGG 1205
Db 667 rProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrProThrGl 687
Qy 1206 ATGGGGCTTGGCCATAGCCATCAGCGCTGAGCTGGAACCTTCTGGCTCCTCTCCGAT 1265
Db 687 yTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLeuProIl 707
Qy 1266 CCATACCTCGGAACTCTCAGCAGCTTGTGTGCTCGCAGCGGCTGAGCAAACTTAT 1325
Db 707 eHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLysLeuIl 727
Qy 1326 CGGAACCGCAACTCTGTGTCTCTCTCGGAATATACACCTCTTCCATCGCTGCTAGG 1385
Db 727 eGlyThrAspAsnSerValValLeuSerArgLysTyRThrSerPheProTrpLeuLeuGl 747
Qy 1386 GTGTGCTGCCAACTGGATCCTCGCGGAGCATCTCTTGTCTACGTCCCGCTCGGCGCTGAA 1445
Db 747 yCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyRValProSerAlaLeuAs 767
Qy 1446 TCCGCGGACACCGCTCTCGGGGCGGTTTGGGCTCTACCGTCCCTCTTCTATCTGCC 1505
Db 767 nProAlaAspAspProSerArgLysGlyArgLeuGlyLeuTyRArgProLeuLeuHisLeuPr 787
Qy 1506 GTTCGGCGGACACCGGCGGACCTCTCTTTACGCGGTCTCCCGTATGTCCTTCTCA 1565
Db 787 oPheArgProThrThrGlyArgThrSerLeuTyRAlaValSerProSerValProSerHi 807
Qy 1566 TCTCGGACCGTGTGCACTTCGCTTCACTCTGCACTGCTGCGATCGGACCAACCG 1620

Db 807 sLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPro 825
RESULT 12
ADX40773
ID ADX40773 standard; protein; 827 AA.
AC ADX40773;
XX 21-APR-2005 (first entry)
XX HBV polymerase protein #16.
XX Immune stimulation; polymerase; enzyme.
XX Hepatitis B virus.
XX WO2005012502-A2.
XX 10-FEB-2005.
XX 29-MAR-2004; 2004WO-US009510.
XX 28-MAR-2003; 2003US-0458026P.
XX (BPIM-) EPIMUNE INC.
XX Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX WPI; 2005-132661/14.
XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
XX response comprises identifying variants of a peptide epitope 8-11 amino
XX acids in length comprising primary anchor residues of the same HLA class
XX I binding motif.
XX Disclosure; Page 380-385; 458pp; English.
XX The invention relates to a method of identifying a candidate peptide
XX epitope which induces an HLA class I CTL response against variants of the
XX peptide epitope, comprising identifying, from a particular antigen of an
XX infectious agent, variants of a peptide epitope comprising primary anchor
XX residues of the same HLA class I binding motif. The method is useful for
XX identifying a candidate peptide epitope, which induces an HLA class I CTL
XX response against variants of the peptide epitope. This sequence
XX represents an HBV polymerase protein used in the scope of the invention.
XX Sequence 827 AA;
SQ
Alignment Scores:
Pred. No.: 5.57e-204 Length: 827
Score: 2678.50 Matches: 504
Percent Similarity: 95.91% Conservative: 12
Best Local Similarity: 93.68% Mismatches: 18
Query Match: 45.34% Indels: 5
DB: Gaps: 1
US-10-761-006A-1 (1-3215) x ADX40773 (1-827)
Qy 19 AGCTCTCTAGATCCAGGGTGAGGG-----CCTATATTTTCTGCTGTGGTCTC 69
Db 291 ThrSerLysArgGlnSerSerSerGlyHisAlaValGluLeuHisProCysTrpTrpLeu 310
Qy 70 CAGTTCGGAACAGTAACCTGTTCCGACTACTGCTCTCCCATATCGTCAATCTTCTC 129
Db 311 GlnPheArgAsnSerLysProCysSerAspTyRcysLeuThrHisIleValAsnLeuLeu 330
Qy 130 GAGGACTGGGAGCCCTCCACCGCAACATGGAGAACACAACTCAGGATTCCTAGACCCCT 189
Db 331 GluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIleProArgTrpPro 350
Qy 190 GCTCGTCTTACAGCGGGGTTTCTGTTGACAGAGTCTCTCAATACCGCAGTCT 249
Db 351 AlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThrThrGluSer 370

QY 250 AGACTCTG-GTGGACTTCTCTCAATTTCTAGGGGAGCACCACGCTTCTCTGGCCAAA 308
Db |||||
QY 371 ArgLeuValValAspPheSerGlnPheSerArgGlySerThr***Val-SerTrpProLy 390
Db |||||
QY 309 ATTGGCAGTCCCAACCTCCAACTACTCACAACCTCTTGTCTCTCCATTTGTCTCTGGCT 368
Db |||||
QY 330 sPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeuSerTrpLe 410
Db |||||
QY 369 ATCGCTGGATGTCTGGGGCGTATTATCATATTCCTTTCATCTCTGCTGTATGCTCA 428
Db |||||
QY 410 uSerLeuAspValSerAlaAlaPheTyHisileProLeuHisProAlaAlaMetProHi 430
Db |||||
QY 429 TCTTCTGTGTGTTCTTGTGACTACCAAGGTATGTTGCCGTTTGTCTCTACTTCCAG 488
Db |||||
QY 430 sLeuLeuValGlySerSerGlyLeuProArgTyValAlaArgLeuSerSerThrSerAr 450
Db |||||
QY 489 GAACATCAACACCAAGCAGCGGCCATGCAAGACCTGCACGACTCTCTCAAGGAAACT 548
Db |||||
QY 450 GAsn***AsnTyGlnHisGlyThrMetGlnAspLeuHisGlnSerCysSerArgAsnLe 470
Db |||||
QY 549 CTACGTTTCCCTCTTGTCTGTATACAAACCTTCGGACGGAACCTGCTATTTATTCCTCA 608
Db |||||
QY 470 uTyValSerLeuLeuLeuLeuTyLysThrPheGlyArgLysLeuHisLeuTySerHi 490
Db |||||
QY 609 TCCCATCATCTGGGCTTCGCAAGATTCCTATGGAGTGGCCCTCAGTCGGTTCTCTCT 668
Db |||||
QY 490 sProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerProPheLeuLe 510
Db |||||
QY 669 GCCTCAGTTTACTAGTGCATTTGTTTCAGTGTCTGTAGGGCTTCCGCCACTGTTTGGC 728
Db |||||
QY 510 uAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHisCysLeuAl 530
Db |||||
QY 729 TTTTCAGTTATATGATGTGTGTATGGGGCGAAGTCTGTACAAACATCTTGAGTCCT 788
Db |||||
QY 530 aPheSerTyMetAspAspValValLeuGlyAlaLysSerValGlnHisLeuGluSerLe 550
Db |||||
QY 789 TTTTACCTCTATACCAATTTCTTGTCTTGTGGGTATACATTTAAACCTTAATAAAC 848
Db |||||
QY 550 uPheThrAlaIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnProHisLysTh 570
Db |||||
QY 849 CAAAGCTGGGGCTACTCCCTAACTTCATGGGATATGTAATTGGAAGTTGGGGTACTTT 908
Db |||||
QY 570 rLysArgTrpGlyTySerLeuAsnPheMetGlyTyValIleGlySerTrpGlyThrLe 590
Db |||||
QY 909 ACCGAGGAAACATATTGTACTAAACTCAAGCAATGTTTTCGAAAACCTGCTGTAATAG 968
Db |||||
QY 590 uProGlnGluHisIleValGlnLysIleLysGlnCysPheArgLysLeuProValAsnAr 610
Db |||||
QY 969 ACCTATTGATGGAAAGTATGCAAGAAATGTGGGTCTTTTGGGCTTGTGGCCCTTT 1028
Db |||||
QY 610 gProIleAspTrpLysValCysGln***IleValGlyLeuLeuGlyPheAlaAlaProPh 630
Db |||||
QY 1029 TACACAATGTGCTATCTGCTCTGATGCTTTATATGATCATGTATACATCTCAACGAGC 1088
Db |||||
QY 630 eThrGlnCysGlyTyProAlaLeuMetProLeuTyAlaCysIleGlnAlaLysGlnAl 650
Db |||||
QY 1089 TTTTCACTTCTCGCCCAACTTACAAGGCTTCTTGTGTAAACAATATCTGAACCTTTACCC 1148
Db |||||
QY 650 aPheThrPheSerProThrTyLysAlaPheLeuCysLysGlnTyLeuHisLeuTyPr 670
Db |||||
QY 1149 CTTGCCCCGGCAAGCTCGGCTCTGCGCAAGTGTGTCGACGCAACCCCACTGGAGT 1208
Db |||||
QY 670 vAlaAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrProThrGlyTr 690
Db |||||
QY 1209 GGGCTGGCATAGGCCATCAGCGCATGCTGGAACCTTCTGGCTCTCTGCCCATCA 1268
Db |||||
QY 690 pGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLeuProIleHi 710
Db |||||
QY 1269 TACTGGGAACCTCTAGCAGCTTGTGTGCTGCGACCGCTCTGGAGCAAACTTATCGG 1328
Db |||||
QY 710 sThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLysLeuIleGl 730
Db |||||

QY 1329 AACGCAACTCTGTTGTTCTCTCTCTCGGAATAACACCTCTCTTCATGGCTGCTAGGTG 1388
Db |||||
QY 730 yThrAspAsnSerValValLeuSerArgLysTyThrSerPheProTrpLeuLeuGlyCy 750
Db |||||
QY 1389 TGCTGCCAACTGGATCTCGCGGAGCTCTTTTGTCTTACGTCCCGTGGCGCTGAATCC 1448
Db |||||
QY 750 sAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyValProSerAlaLeuAsnPr 770
Db |||||
QY 1449 CGCGAGACACCGCTCTCGGGCCGTTTGGGCTCTACCGTCCCTTCTTCATCTGCCCTT 1508
Db |||||
QY 770 oAlaAspAspProSerArgGlyArgLeuGlyLeuTyArgProLeuLeuHisLeuProPh 790
Db |||||
QY 1509 CGCGCCGACCAACGCGGCGACCTCTCTTACGCGTCTCCCGTATGCTTCTTCATCT 1568
Db |||||
QY 790 eArgProThrThrGlyArgThrSerLeuTyAlaValSerProSerValProSerHisLe 810
Db |||||
QY 1569 GCCGAGCGGTGTGCATCTTCGCTTCACCTCTGCACGTCCGATGGAGACACCG 1620
Db |||||
QY 810 uProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPro 827
Db |||||
RESULT 13
AAE04708
ID AAE04708 standard; protein; 842 AA.
XX
AC AAE04708;
XX
DT 04-SEP-2001 (first entry)
XX
DE Hepatitis B virus FRI strain genotype G HBpol protein.
XX
KW HBV genotype G; precore; HBpol; polymerase; envelope protein; preS1;
KW preS2; surface antigen; HBeAg; HBx protein; vaccine; liver disease;
KW hepatitis; liver cancer; HBeAg; core antigen.
XX
OS Hepatitis B virus.
XX
PN WO200138498-A2.
XX
PD 31-MAY-2001.
XX
PF 21-NOV-2000; 2000WO-US032108.
XX
PR 24-NOV-1999; 99US-0167206P.
XX
PA (PHAR-) PHARMASSET INC.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Stuyver L, Schinazi R, De Gendt S, Van Geyt C, Zoulim F, Fried M;
PI Roseau R;
XX
XX WPI; 2001-367676/38.
DR N-PSDB; AAD09091.
XX
PT Novel hepatitis B virus genotype G, nucleic acids encoding virus,
PT polypeptides encoded by nucleic acids, useful for preparing vaccine to
PT treat or prevent the hepatitis B virus genotype G infection in a subject.
XX
XX Example; Fig 6; 84pp; English.
PS
CC The present invention relates to hepatitis B virus (HBV) strain FRI,
CC genotype G DNA encoding PreCore/Core protein, HBpol, envelope (PreS1,
CC PreS2 and surface antigen HBeAg) and HBx proteins. HBV genotype G nucleic
CC acids and polypeptides are useful for diagnosing, prognosing and treating
CC infections caused by HBV genotype G. They can be used in a vaccine to
CC treat or prevent HBV genotype G infection. The HBV genotype G derived
CC nucleic acids and antibodies are useful for detecting HBV genotype G in a
CC sample or diagnosis of HBV genotype G infection. The presence of HBV
CC genotype G statistically correlates with the presence of liver damage
CC and/or liver cancer in the subject. The HBV genotype G core insert
CC peptide encoding nucleic acid is useful for designing monitoring assays
CC to study and predict the evolution of anti-HBe and anti-HBc antibodies
CC and HBeAg (genotype G e antigen) in patients infected with HBV. The
CC antibodies or antigens of HBV genotype G are useful for identifying a

QY 1500 TCTGCGCTTCGGCCGACACCGGGCGCACCTCTCTTTACGCGGTCTCCCGTATGTC 1559
 Db 802 gLeuProPheLeuProThrThrGlyArgThrSerLeuTyzAlaValSerProSerValPr 822
 QY 1560 TTCTCATCTCCGACCGCTGTCGCTTCACCTCTGCGCTGCGATGGAGACCC 1619
 Db 822 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValThrTrpLysProPr 842
 QY 1620 G 1620
 Db 842 o 842
 RESULT 15
 ID ADX40761 standard; protein; 832 AA.
 XX AC ADX40761;
 XX DT 21-APR-2005 (first entry)
 XX DE HBV polymerase protein #4.
 XX KW Immune stimulation; polymerase; enzyme.
 XX OS Hepatitis B virus.
 XX FN WO2005012502-A2.
 XX PD 10-FEB-2005.
 XX PF 29-MAR-2004; 2004WO-US009510.
 XX PR 28-MAR-2003; 2003US-0458026P.
 XX PA (EPIM-) EPIMUNE INC.
 XX PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
 XX WPI; 2005-132661/14.
 XX PT Identifying a candidate peptide epitope, which induces a HLA class I CTL
 PT response comprises identifying variants of a peptide epitope 8-11 amino
 PT acids in length comprising primary anchor residues of the same HLA class
 PT I binding motif.
 XX PS Disclosure; Page 380-385; 458pp; English.
 CC The invention relates to a method of identifying a candidate peptide
 CC epitope which induces an HLA class I CTL response against variants of the
 CC peptide epitope, comprising identifying, from a particular antigen of an
 CC infectious agent, variants of a peptide epitope comprising primary anchor
 CC residues of the same HLA class I binding motif. The method is useful for
 CC identifying a candidate peptide epitope, which induces an HLA class I CTL
 CC response against variants of the peptide epitope. This sequence
 CC represents an HBV polymerase protein used in the scope of the invention.
 XX SQ Sequence 832 AA;
 Alignment Scores:
 Pred. No.: 4.97e-202 Length: 832
 Score: 2654.00 Matches: 492
 Percent Similarity: 95.56% Conservative: 25
 Best Local Similarity: 90.94% Mismatches: 23
 Query Match: 44.93% Indels: 2
 DB: 9 Gaps: 0
 US-10-761-006A-1 (1-3215) x ADX40761 (1-832)
 QY 1 CTCCACCAACATTCACCAAGCTCTGCTAGATCCCGGCTGAGGGCTATATTTCTCTGC 60
 Db 293 LeuHisAsnPheProProAsnSerAlaArgSerGlnGlyGluArgProValPheProCys 312
 QY 61 TGGTGGCTCCAGTTCCTCCGGAACAGTAACCCCTGTTCCGACTACTCCCTCTCCCATATCGTC 120

313 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerHisIleVal 332
 QY 121 AATCTTCTCGAGACATGGGACCCCTCACCGAACAATGGAGAACACAAACATCAGATTTCCT 180
 Db 333 AsnLeuLeuGlnAspTrpGlyProCysThrGluHisGlyGluHisIleArgIlePro 352
 QY 181 AGGACCCCTGCTCGTGTATACAGCGGGGTTTTCTCGTTGACAAAGAAATCCTCAAAATACC 240
 Db 353 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 372
 QY 241 GCAGAGCTTAGACTCTG-GTGACTTCTCTCAATTTTCTAGGGGAGACCCACAGTGTTC 299
 Db 373 AlaGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnHisArgVal-Se 392
 QY 300 CTGGCCAAATTCGACAGTCCCCAACCTCCAATCACCTCACCAACCTCTGTGCTCCCAATTT 359
 Db 392 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 412
 QY 360 GTCTCGGCTATCGCTGGATGTGTCTCGGGGTTTTATCATATTCTCTCTCATCTGTCTGC 419
 Db 412 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaSe 432
 QY 420 TATGCTCATCTCTCTGTTGTTGTTCTCTGACTACCAAGTATGTTGCCGCTTGTCTCCTC 479
 Db 432 rMetProHisLeuLeuValGlySerThrGlyLeuSerArgTyrValAlaArgValSerSe 452
 QY 480 TACTTCCAGGAACATCAACCCAGCACGGGGCCATGCAAGACCTGCACCACTCTCTGCTC 539
 Db 452 rAsnSerArgIlePheAsnHisGlnArgGlyThrMetGlnAsnLeuHisAspTyrCysSe 472
 QY 540 AAGGAAACTCTACGTTTCCCTCTGTTGTTGTGTGTAACAAACCTTCGGACGGAAACTGCACTT 599
 Db 472 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrGlnThrPheGlyArgLysLeuHisLe 492
 QY 600 GTATTCCTCCATCCCATCATCTGCTGGCTTTCGCAAGATCTCTATGGAGTGGGCTCAGTCC 659
 Db 492 uTyrSerHisProIleLeuLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 512
 QY 660 GTTCTCTCGGCTCAGTTTACTAGTGCATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 719
 Db 512 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi 532
 QY 720 CTGTTTGGCTTTCAGTTTATATGATGATGTGTATTTGGGGCGGAAGTCTGTACAACATCT 779
 Db 532 sCysLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLe 552
 QY 780 TGAGTCCCTTTTACCTCTATATACCAATTTTCTTTGCTTTGGGTATATACATTAACCC 839
 Db 552 uGluSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 572
 QY 840 TAATAAAACCAACGTTGGGGCTACTCCCTTAACTTCATCGGATATGTAATTTGGAAGTGTG 899
 Db 572 oAsnLysThrLysArgTrpGlyTrpSerLeuHisPheMetGlyTyrValIleGlySerTy 592
 QY 900 GGGTACTTTTACCCGAGAACATATTTGTACTTAAACCTCAAGCAATGTTTTCGAAACATGCC 959
 Db 592 rGlySerLeuProGlnAspHisIleIleGlnLysIleLysGluCysPheArgLysLeuPr 612
 QY 960 TGTAAATAGACCTATTGATTGGAAAGTATGTCAAGAAATTTGTGGGTCTTTTGGCTTTGC 1019
 Db 612 oIleAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 632
 QY 1020 TGCCCTTTTACACAAATGTGGCTATCTGCTGCTGTGATGCCCTTTATATCATGTATACAAATC 1079
 Db 632 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 652
 QY 1080 TAAGCAGGCTTTTCACTTTCTCGCCAACTTCAAGAGGCTTTCTGTGTAAACAATATCTGAA 1139
 Db 652 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 672
 QY 1140 CTTTACCCGCTGTCGGCGCAACGGTCCGCTCTCTGCAAGTGTGTGCTACGCAACCC 1199

```

Db      672 nLeuTyPrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 692
QY      1200 CACTGGATGGGCTTGGCCATAGGCCATAGCGCATGGCTGGAACCTTTCTGGCTCTCT 1259
Db      692 oThrGlyTrpGlyLeuValMetGlyHisGlnArgMetArgGlyThrPheLeuAlaProLe 712
QY      1260 GCCGATCCATAGTGGGAACTCCTAGCAGCTGTGTTGCTCGCAGCCGGTCTGGAGCAAA 1319
Db      712 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaAs 732
QY      1320 ACTTATCGGAACCGCAACTCTGTTGTCTCTCTCGGAAATACACCTCTTTCCATGGCT 1379
Db      732 nIleLeuGlyThrAspAsnSerValIleuSerArgIysTyThrSerPheProTribLe 752
QY      1380 GCTAGGGTGTGTGTCAACTGGATCCTGCGGGGACGTCTTTGTGTACGTCCCGTCGGC 1439
Db      752 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyValProSerAl 772
QY      1440 GCTGAATCCCGGGACGACCGCTCTCGGGCGGTTGGGGCTCTACCGTCCGCTCTTCA 1499
Db      772 aLeuAsnProAlaAspAspProSerArgGlyIlePheArgProLeuLeuAr 792
QY      1500 TCTGCGCTTCGGGCGGACCGCGGCGCACCTCTCTTTACGGGTCTCCCGTATGTGCC 1559
Db      792 gLeuProPheArgProThrThrGlyArgThrSerLeuTyAlaAspSerProSerValPr 812
QY      1560 TTCTCATCTGCGGACCGGTGTGCACTTCGCTTCACCTCTGCAACGTTCGATGGAGACCACC 1619
Db      812 oSerHisLeuProValArgValHisPheAlaSerProLeuHisValAlaTTrpArgProPr 832
QY      1620 G 1620
Db      832 o 832

```

Search completed: December 1, 2005, 23:35:15
 Job time : 520.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: December 27, 2005, 20:42:07 ; Search time 3.91 Seconds
(without alignments)
3395.886 Million cell updates/sec

Title: US-10-761-006A-1_COPY_527_595
Perfect score: 139
Sequence: 1 ACAGCTCTGCTCAGGAAA.....AACCTTCGGACGGAARCTGC 69

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-O=/cgn2_1/USPTO_epool_p/US10761006/runat_27122005_192901_15763/app_query.fasta_1.1102
-DB_PIR -QFMT-fastan -SUFFIX=rpr -MINMATCH=0.1 -LOPPEXT=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORW=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USBR=US10761006 @CNG 1 1 77 @runat_27122005_192901_15763 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR 80:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	133	95.7	226	JQ2054	surface antigen -
2	130	93.5	226	JQ1577	major surface anti
3	130	93.5	226	JQ1577	major surface anti
4	130	93.5	226	JQ2052	surface antigen -
5	130	93.5	226	JQ2053	surface antigen -
6	130	93.5	226	JQ2050	surface antigen -
7	130	93.5	226	JQ2058	surface antigen -
8	130	93.5	226	JQ2055	surface antigen -
9	130	93.5	226	JQ2078	surface antigen -
10	130	93.5	226	JQ2046	surface antigen -
11	130	93.5	226	JQ2045	surface antigen -
12	130	93.5	226	JQ2056	surface antigen -
13	130	93.5	389	JQ1576	large surface anti
14	130	93.5	400	JQ1575	major surface anti

15	130	93.5	400	1	SAVLKS	large surface anti
16	130	93.5	400	1	SAVLVD	large surface anti
17	127	91.4	226	1	JQ1574	major surface anti
18	127	91.4	226	1	SAVLAD	major surface anti
19	127	91.4	226	2	JQ2101	surface antigen -
20	125	89.9	389	2	S20749	surface antigen -
21	124	89.2	226	1	JQ1570	major surface anti
22	124	89.2	226	1	JQ1571	major surface anti
23	124	89.2	226	2	JQ2058	surface antigen -
24	124	89.2	226	2	JQ2075	surface antigen -
25	124	89.2	226	2	JQ2057	surface antigen -
26	124	89.2	226	2	JQ2077	surface antigen -
27	124	89.2	226	2	JQ2076	surface antigen -
28	124	89.2	226	2	JQ2061	surface antigen -
29	124	89.2	226	2	JQ2081	surface antigen -
30	124	89.2	226	2	JQ2060	surface antigen -
31	124	89.2	226	2	JQ2079	surface antigen -
32	124	89.2	226	2	JQ2083	surface antigen -
33	124	89.2	226	2	JQ2062	surface antigen -
34	124	89.2	389	1	SAVLAI	large surface anti
35	124	89.2	389	1	SAVLBI	large surface anti
36	124	89.2	389	1	SAVLJ2	large surface anti
37	124	89.2	389	1	SAVLJ3	large surface anti
38	124	89.2	389	2	S20745	surface antigen -
39	124	89.2	389	2	S20753	surface antigen -
40	124	89.2	389	2	S41871	surface antigen -
41	123	88.5	226	2	JQ2116	surface antigen -
42	122	87.8	226	2	JQ2047	surface antigen -
43	122	87.8	226	2	JQ2051	surface antigen -
44	122	87.8	384	2	T13469	large surface anti
45	122	87.8	389	1	SAVLJ1	large surface anti

ALIGNMENTS

RESULT 1

JQ2054 surface antigen - hepatitis B virus (subtype indet, strain Tar)

C/Species: hepatitis B virus, HBV

C/Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 31-Dec-2004

C/Accession: JQ2054; PQ0571

R/Norder, H.; Hammam, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnus

J. Gen. Virol. 74, 1341-1348, 1993

A/Title: Genetic relatedness of hepatitis B virus strains of diverse geographical origi

A/Reference number: JQ2044; MUID:93329382; PMID:8336122

A/Contents: genogroup A

A/Accession: JQ2054

A/Molecule type: DNA

A/Residues: 1-226 <NOR>

A/Cross-references: UNIPROT:Q8JWP7; UNIPROT:Q9Q3D4; UNIPROT:Q8JXH6; UNIPROT:O41759; UNI

PROT:O41755; UNIPROT:O41746; UNIPROT:O90663; UNIPROT:O41744; UNIPROT:Q9JEX6; UNIPROT:O4

JXH0; UNIPROT:Q8JXH3; UNIPROT:Q91C42; UNIPROT:Q8JLX2; UNIPROT:Q91L67; UNIPROT:Q9W7M6; U

R/Norder, H.; Courouce, A.M.; Magnus, L.O.

J. Gen. Virol. 73, 3141-3145, 1992

A/Title: Molecular basis of hepatitis B virus serotype variations within the four major

A/Reference number: PQ0453; MUID:93107848; PMID:1469353

A/Accession: PQ0571

A/Molecule type: DNA

A/Residues: 101-180 <NO2>

A/Cross-references: UNIPARC:UPI0000178434

C/Genetics:

A/Gene: S

C/Keywords: surface antigen

Alignment Scores:	7.82e-12	Length:	226
Pred. No.:	133.00	Matches:	22
Score:	95.65%	Conservative:	0
Percent Similarity:	95.65%	Mismatches:	1
Best Local Similarity:	95.65%	Indels:	0
Query Match:	95.68%	Gaps:	0
DB:	2		

US-10-761-006A-1_COPY_527_595 (1-69) x JQ2054 (1-226)

```
QY 1 ACGACTCCTGCTCAAGGAACTCTACGTTTCCTCTTGTTGCTGTACAAAACCTTCGGAC 60
Db 125 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProSerAsp 144
QY 61 GGAAACTGC 69
Db 145 GlyAsnCys 147

RESULT 2
JQ1577
major surface antigen - hepatitis B virus (subtype ayw1)
C:Species: hepatitis B virus, HBV
A:Note: host Homo sapiens (man)
C:Date: 31-Dec-1993 #sequence revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: JQ1577; JQ2049; PQ0573
R:Order: H.; Hammas, B.; Loefdaahl, S.; Courouce, A.M.; Magnius, L.O.
J. Gen. Virol. 73, 1201-1208, 1992
A:Title: Comparison of the amino acid sequences of nine different serotypes of hepatitis
A:Reference number: JQ1570; MUID:92268879; PMID:1588323
A:Accession: JQ1577
A:Molecule type: DNA
A:Residues: 1-226 <NOR>
A:Cross-references: UNIPROT:Q69589; UNIPARC:UPI00000F4DF5; GB:X75669; NID:9416076; PIDN:
A:Experimental source: subtype ayw1, strain CNTS-5
R:Order: H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnius,
J. Gen. Virol. 74, 1341-1348, 1993
A:Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A:Reference number: JQ2044; MUID:93329382; PMID:8336122
A:Contents: genogroup A
A:Accession: JQ2049
A:Molecule type: DNA
A:Residues: 1-226 <NO2>
A:Cross-references: UNIPARC:UPI00000F4DF5
A:Experimental source: subtype ayw1, strain Mam
R:Order: H.; Courouce, A.M.; Magnius, L.O.
J. Gen. Virol. 73, 3141-3145, 1992
A:Title: Molecular basis of hepatitis B virus serotype variations within the four major
A:Reference number: PQ0453; MUID:93107848; PMID:1469353
A:Accession: PQ0573
A:Molecule type: DNA
A:Residues: 101-180 <NO3>
A:Cross-references: UNIPARC:UPI000017496F
A:Experimental source: subtype ayw1, strain Mam
C:Genetics:
A:Gene: S
C:Superfamily: hepatitis B virus surface antigen
C:Keywords: glycoprotein, surface antigen
F:3/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.: 2,23e-11 Length: 226
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0
DB: 1 Gaps: 0

US-10-761-006a-1_COPY_527_595 (1-69) x JQ1577 (1-226)

QY 1 ACGACTCCTGCTCAAGGAACTCTACGTTTCCTCTTGTTGCTGTACAAAACCTTCGGAC 60
Db 125 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProThrAsp 144
QY 61 GGAAACTGC 69
Db 145 GlyAsnCys 147

RESULT 3
SAVLHV
major surface antigen - hepatitis B virus
C:Species: hepatitis B virus, HBV
A:Note: host Homo sapiens (man)
```

```
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C:Accession: JQ2093
R:Rivkina, M.B.; Lunin, V.G.; Mahov, A.M.; Tikchonenko, T.I.; Kukain, R.A.
Gene 64, 285-296, 1988
A:Title: Nucleotide sequence of integrated hepatitis B virus DNA and human flanking regi
A:Reference number: JQ2093; MUID:86297159; PMID:2841200
A:Accession: JQ2093
A:Molecule type: DNA
A:Residues: 1-226 <RIV>
A:Cross-references: UNIPROT:P31873; UNIPARC:UPI00001389B6; GB:M21030; NID:9329702; PIDN:
C:Genetics:
A:Gene: S
C:Superfamily: hepatitis B virus surface antigen
C:Keywords: glycoprotein, surface antigen
F:3/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.: 2,23e-11 Length: 226
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0
DB: 1 Gaps: 0

US-10-761-006a-1_COPY_527_595 (1-69) x SAVLHV (1-226)

QY 1 ACGACTCCTGCTCAAGGAACTCTACGTTTCCTCTTGTTGCTGTACAAAACCTTCGGAC 60
Db 125 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProThrAsp 144
QY 61 GGAAACTGC 69
Db 145 GlyAsnCys 147

RESULT 4
JQ2052
surface antigen - hepatitis B virus (subtype adw2, strain 8903/86)
C:Species: hepatitis B virus, HBV
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 09-Jul-2004
C:Accession: JQ2052
R:Order: H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnius,
J. Gen. Virol. 74, 1341-1348, 1993
A:Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A:Reference number: JQ2044; MUID:93329382; PMID:8336122
A:Contents: genogroup A
A:Accession: JQ2052
A:Molecule type: DNA
A:Residues: 1-226 <NOR>
A:Cross-references: UNIPROT:Q9W7W6; UNIPARC:UPI00000FA191
C:Genetics:
A:Gene: S
C:Superfamily: hepatitis B virus surface antigen
C:Keywords: surface antigen

Alignment Scores:
Pred. No.: 2,23e-11 Length: 226
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0
DB: 2 Gaps: 0

US-10-761-006a-1_COPY_527_595 (1-69) x JQ2052 (1-226)

QY 1 ACGACTCCTGCTCAAGGAACTCTACGTTTCCTCTTGTTGCTGTACAAAACCTTCGGAC 60
Db 125 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProThrAsp 144
QY 61 GGAAACTGC 69
Db 145 GlyAsnCys 147

RESULT 5
```

Qy 61 GGAAACTGC 69
| | | | |
Db 145 GlyAsnCys 147

RESULT 7

JQ2048
surface antigen - hepatitis B virus (subtype ayw1, strain Bat)
C;Species: hepatitis B virus, HBV
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 31-Dec-2004
C;Accession: JQ2048
R;Order, H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnius J. Gen. Virol. 74, 1341-1348, 1993
A;Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin.
A;Reference number: JQ2044; MUID:93329382; PMID:8336122
A;Contents: genogroup A
A;Accession: JQ2048
A;Molecule type: DNA
A;Residues: 1-226 <NOR>
A;Cross-references: UNIPROT:Q9DLM8; UNIPROT:Q9AWJ7; UNIPROT:Q9Q3D4; UNIPROT:Q9YVD5; UNIPROT:O41755; UNIPROT:O41746; UNIPROT:O90663; UNIPROT:O41744; UNIPROT:O9WRK7; UNIPROT:Q8UXH0; UNIPROT:Q8UXH3; UNIPROT:Q91C42; UNIPROT:Q8JLX2; UNIPROT:Q91L67; UCRK; UNIPROT:Q8UXH0; UNIPROT:Q8UXH3; UNIPROT:Q91C42; UNIPROT:Q8JLX2; UNIPROT:Q91L67; U

C;Genetics:
A;Gene: S
C;Keywords: surface antigen

Alignment Scores:	2.23e-11	Length:	226
Pred. No.:	130.00	Matches:	21
Score:	95.65%	Conservative:	1
Percent Similarity:	91.30%	Mismatches:	1
Best Local Similarity:	93.53%	Indels:	0
Query Match:	2	Gaps:	0
DB:			

US-10-761-006A-1_COPY_527_595 (1-69) x JQ2048 (1-226)

Qy 1 ACGACTCTGCTCAAGAAACTACCTTTCCTTTGGCTGTACAAAACCTTCGGAC 60
| | | | |
Db 125 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrIysProThrAsp 144

Qy 61 GGAAACTGC 69
| | | | |
Db 145 GlyAsnCys 147

RESULT 8

JQ2055
surface antigen - hepatitis B virus (subtype adw2, strain Bl-85)
C;Species: hepatitis B virus, HBV
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 31-Dec-2004
C;Accession: JQ2055
R;Order, H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnius J. Gen. Virol. 74, 1341-1348, 1993
A;Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin.
A;Reference number: JQ2044; MUID:93329382; PMID:8336122
A;Contents: genogroup A
A;Accession: JQ2055
A;Molecule type: DNA
A;Residues: 1-226 <NOR>
A;Cross-references: UNIPROT:Q81186; UNIPARC:UPI0000017842E
C;Genetics:
A;Gene: S
C;Keywords: surface antigen

Alignment Scores:	2.23e-11	Length:	226
Pred. No.:	130.00	Matches:	21
Score:	95.65%	Conservative:	1
Percent Similarity:	91.30%	Mismatches:	1
Best Local Similarity:	93.53%	Indels:	0
Query Match:	2	Gaps:	0
DB:			

US-10-761-006A-1_COPY_527_595 (1-69) x JQ2055 (1-226)

```
Db      125 ThrThrProAlaGlnGlyAasnSerMetPheProSerCysCysCysThrLysProThrAasp 144
      |||
QY      61 GGAAACTGC 69
      |||
Db      145 GlyAasnCys 147

RESULT 9
JQ2078
surface antigen - hepatitis B virus (subtype ayw, strain aywmnt)
C:Species: hepatitis B virus, HBV
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 31-Dec-2004
C:Accession: JQ2078
R:Norder, H.; Hammass, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnus,
J. Gen. Virol. 74, 1341-1348, 1993
A:Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A:Reference number: JQ2044; MUID:93329382; PMID:8336122
A:Contents: genogroup D
A:Accession: JQ2078
A:Molecule type: DNA
A:Residues: 1-226 <NOR>
A:Cross-references: UNIPROT:Q67886; UNIPARC:UPI0000178428
C:Genetics:
A:Gene: S
C:Keywords: surface antigen

Alignment Scores:
Pred. No.:      2,23e-11      Length:      226
Score:          130.00      Matches:      21
Percent Similarity: 95.65%      Conservative: 1
Best Local Similarity: 91.30%      Mismatches: 1
Query Match:    93.53%      Indels:      0
DB:             2          Gaps:      0

US-10-761-006a-1_COPY_527_595 (1-69) x JQ2078 (1-226)

QY      1  ACGACTCTGCTCAAGGAAACTCTACGTTTCCTTCTTGCTGTACAAAACCTTCGGAC 60
      |||
Db      125 ThrThrProAlaGlnGlyAasnSerMetTyrProSerCysCysCysThrLysProSerAasp 144
      |||
QY      61 GGAAACTGC 69
      |||
Db      145 GlyAasnCys 147

RESULT 10
JQ2046
surface antigen - hepatitis B virus (subtype adw2, strain Wil)
C:Species: hepatitis B virus, HBV
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 31-Dec-2004
C:Accession: JQ2046
R:Norder, H.; Hammass, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnus,
J. Gen. Virol. 74, 1341-1348, 1993
A:Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A:Reference number: JQ2044; MUID:93329382; PMID:8336122
A:Contents: genogroup A
A:Accession: JQ2046
A:Molecule type: DNA
A:Residues: 1-226 <NOR>
A:Cross-references: UNIPROT:Q91HD6; UNIPROT:Q9Q3C9; UNIPROT:Q72529; UNIPROT:Q91EH7; UNIP
PROT:Q9Q3D4; UNIPROT:Q39886; UNIPROT:Q41754; UNIPROT:Q42029; UNIPROT:Q91
1EH8; UNIPROT:Q8JWP5; UNIPROT:Q9WRK7; UNIPROT:Q8JXH0; UNIPROT:Q91C42; UNIPROT:Q8JLX2; UN
C:Genetics:
A:Gene: S
C:Keywords: surface antigen

Alignment Scores:
Pred. No.:      2,23e-11      Length:      226
Score:          130.00      Matches:      21
Percent Similarity: 95.65%      Conservative: 1
Best Local Similarity: 91.30%      Mismatches: 1
Query Match:    93.53%      Indels:      0
DB:             2          Gaps:      0
```

```
US-10-761-006a-1_COPY_527_595 (1-69) x JQ2046 (1-226)

QY      1  ACGACTCTGCTCAAGGAAACTCTACGTTTCCTTCTTGCTGTACAAAACCTTCGGAC 60
      |||
Db      125 ThrThrProAlaGlnGlyAasnSerMetPheProSerCysCysCysThrLysProThrAasp 144
      |||
QY      61 GGAAACTGC 69
      |||
Db      145 GlyAasnCys 147

RESULT 11
JQ2045
surface antigen - hepatitis B virus (subtype adw2, strain Len)
C:Species: hepatitis B virus, HBV
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 31-Dec-2004
C:Accession: JQ2045
R:Norder, H.; Hammass, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnus,
J. Gen. Virol. 74, 1341-1348, 1993
A:Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A:Reference number: JQ2044; MUID:93329382; PMID:8336122
A:Contents: genogroup A
A:Accession: JQ2045
A:Molecule type: DNA
A:Residues: 1-226 <NOR>
A:Cross-references: UNIPROT:Q9EGU1; UNIPROT:Q91EH7; UNIPROT:Q8V1E9; UNIPROT:Q9JE50; UNIP
PROT:Q39886; UNIPROT:Q41754; UNIPROT:Q42029; UNIPROT:Q91YD5; UNIPROT:Q41
1EH8; UNIPROT:Q8JWP5; UNIPROT:Q9WRK7; UNIPROT:Q8JXH0; UNIPROT:Q91C42; UNIPROT:Q8JLX2; UN
C:Genetics:
A:Gene: S
C:Keywords: surface antigen

Alignment Scores:
Pred. No.:      2,23e-11      Length:      226
Score:          130.00      Matches:      21
Percent Similarity: 95.65%      Conservative: 1
Best Local Similarity: 91.30%      Mismatches: 1
Query Match:    93.53%      Indels:      0
DB:             2          Gaps:      0

US-10-761-006a-1_COPY_527_595 (1-69) x JQ2045 (1-226)

QY      1  ACGACTCTGCTCAAGGAAACTCTACGTTTCCTTCTTGCTGTACAAAACCTTCGGAC 60
      |||
Db      125 ThrThrProAlaGlnGlyAasnSerMetPheProSerCysCysCysThrLysProThrAasp 144
      |||
QY      61 GGAAACTGC 69
      |||
Db      145 GlyAasnCys 147

RESULT 12
JQ2056
surface antigen - hepatitis B virus (subtype adw2, strain Bl-89)
C:Species: hepatitis B virus, HBV
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 31-Dec-2004
C:Accession: JQ2056
R:Norder, H.; Hammass, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnus,
J. Gen. Virol. 74, 1341-1348, 1993
A:Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A:Reference number: JQ2044; MUID:93329382; PMID:8336122
A:Contents: genogroup A
A:Accession: JQ2056
A:Molecule type: DNA
A:Residues: 1-226 <NOR>
A:Cross-references: UNIPROT:Q81185; UNIPROT:Q81184; UNIPARC:UPI0000178449
C:Genetics:
A:Gene: S
C:Keywords: surface antigen

Alignment Scores:
Pred. No.:      2,23e-11      Length:      226
Score:          130.00      Matches:      21
Percent Similarity: 95.65%      Conservative: 1
Best Local Similarity: 91.30%
```

Query Match: 93.53% Indels: 0
DB: 2 Gaps: 0

US-10-761-006a-1_COPY_527_595 (1-69) x JQ2056 (1-226)

QY 1 ACGACTCCTGCTCAAGAACTCTACGTTTCCTCTTGTGCTGTACAAACCTTCGGAC 60
|||||
125 ThrThrProAlaGlnGlyAsnSerMetTyrProSerCysCysThrLysProSerAsp 144
|||||

QY 61 GGAAACTGCG 69
|||||
DB 145 GlyAsnCys 147

RESULT 13
SAVLVE
large surface antigen - hepatitis B virus (subtype adw)
N;Contains: major surface antigen; middle surface antigen
C;Species: hepatitis B virus, HBV
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C;Accession: A93460; JQ2044; A03706
R;Ono, Y.; Onda, H.; Saeada, R.; Igarashi, K.; Sugino, Y.; Nishioka, K.
Nucleic Acids Res. 11, 1747-1757, 1983
A;Title: The complete nucleotide sequences of the cloned hepatitis B virus DNA; subtype
A;Reference number: A93460; MUID:83168919; PMID:6300776
A;Accession: A93460
A;Molecule type: DNA
A;Residues: 1-389 <ON>
A;Cross-references: UNIPROT:P03142; UNIPARC:UPI00001389C4; GB:V00866; GB:J02201
R;Norder, H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnus,
J. Gen. Virol. 74, 1341-1348, 1993
A;Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A;Reference number: JQ2044; MUID:93329382; PMID:8336122
A;Accession: JQ2044
A;Molecule type: DNA
A;Residues: 164-389 <NOR>
A;Cross-references: UNIPARC:UPI0000156BEO
A;Experimental source: genogroup A, subtype adw2, strain PHBV933
C;Genetics:

A;Gene: pre-S1/pre-S2/S
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: glycoprotein; surface antigen
F;109-389/Product: middle surface antigen (gene pre-S2/S) #status predicted <DSA>
F;164-389/Product: major surface antigen (gene S) #status predicted <MSA>
F;4,26,112,166/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.: 2,01e-11 Length: 389
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0
DB: 1 Gaps: 0

US-10-761-006a-1_COPY_527_595 (1-69) x SAVLVE (1-389)

QY 1 ACGACTCCTGCTCAAGAACTCTACGTTTCCTCTTGTGCTGTACAAACCTTCGGAC 60
|||||
288 ThrThrProAlaGlnGlyAsnSerLysPheProSerCysCysThrLysProThrAsp 307
|||||

QY 61 GGAAACTGCG 69
|||||
DB 308 GlyAsnCys 310

RESULT 14
JQ1575
major surface antigen - hepatitis B virus
N;Alternate names: envelope protein; HBs antigen
N;Contains: surface antigen pre-S1 (large envelope protein); surface antigen pre-S2 (mid
C;Species: hepatitis B virus, HBV
A;Variety: subtype adw2
C;Date: 31-Dec-1993 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C;Accession: S47411; JQ1575
R;Plucienniczak, A.

submitted to the EMBL Data Library, August 1994
A;Description: Molecular cloning and sequencing of two complete genomes of polish isolat
A;Reference number: S47404
A;Accession: S47411
A;Molecule type: DNA
A;Residues: 1-400 <PLU>
A;Cross-references: UNIPROT:Q67896; UNIPARC:UPI00000F5BAF; EMBL:Z35717; NID:9527440; PI
R;Norder, H.; Hammas, B.; Loeferl, S.; Courouce, A.M.; Magnus, L.O.
J. Gen. Virol. 73, 1201-1208, 1992
A;Title: Comparison of the amino acid sequences of nine different serotypes of hepatiti
A;Reference number: JQ1570; MUID:92268879; PMID:1588323
A;Accession: JQ1575
A;Molecule type: DNA
A;Residues: 175-400 <NOR>
A;Cross-references: UNIPARC:UPI00000002B4; GB:X75666; NID:G416074; PIDN:CAAS3362.1; PID
A;Experimental source: subtype adw2, strain P6
C;Genetics:

A;Gene: S
A;Introns: 122/3
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: glycoprotein; surface antigen
F;1-400/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PSI
F;1-119/Domain: pre-S1 domain #status predicted <PRE1>
F;120-400/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <
F;120-174/Domain: pre-S2 domain #status predicted <PRE2>
F;175-400/Product: surface antigen S (small envelope protein) #status predicted <PSD>
F;177/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.: 2e-11 Length: 400
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0
DB: 1 Gaps: 0

US-10-761-006a-1_COPY_527_595 (1-69) x JQ1575 (1-400)

QY 1 ACGACTCCTGCTCAAGAACTCTACGTTTCCTCTTGTGCTGTACAAACCTTCGGAC 60
|||||
299 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProThrAsp 318
|||||

Db 61 GGAAACTGCG 69
|||||
Db 319 GlyAsnCys 321

RESULT 15
SAVLVS
large surface antigen - hepatitis B virus (subtype adw, strain 991)
N;Contains: major surface antigen; middle surface antigen
C;Species: hepatitis B virus, HBV
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C;Accession: S10383
R;Koechel, H.G.; Schueler, A.; Lottmann, S.; Thomssen, R.
submitted to the EMBL Data Library, February 1990
A;Reference number: S10380
A;Accession: S10383
A;Molecule type: DNA
A;Residues: 1-400 <KOE>
A;Cross-references: UNIPROT:P17101; UNIPARC:UPI00001389B8; EMBL:X51970; NID:G1155012; P
C;Genetics:

A;Gene: pre-S1/pre-S2/S
A;Introns: 123/2
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: glycoprotein; surface antigen
F;120-400/Product: middle surface antigen (gene pre-2/S) #status predicted <DSA>
F;175-400/Product: major surface antigen (gene S) #status predicted <MSA>
F;15,123,177/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.: 2e-11 Length: 400
Score: 130.00 Matches: 21

Percent Similarity: 95.65% Conservative: 1
 Best Local Similarity: 91.30% Mismatches: 1
 Query Match: 93.53% Indels: 0
 DB: 1 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x SAVLKS (1-400)

QY 1 ACGACTCCTGCTCAAGGAACTCTACGTTCCCTCTTGTGTGCTGTACAAAACCTTCGGAC 60
 |||||
 Db 299 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProThrAsp 318
 |||||
 QY 61 GGAAGCTGC 69
 |||||
 Db 319 GlyAsnCys 321
 |||||

Search completed: December 27, 2005, 21:01:54
 Job time : 4.91 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 27, 2005, 20:41:12 ; Search time 23.046 Seconds
(without alignments)
4224.720 Million cell updates/sec

Title: US-10-761-006A-1_COPY_527_595

Perfect score: 139

Sequence: 1 ACAGCTCTGCTCAGGAAA.....AACCTTCGACGGAACCTGC 69

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DBV=xlp
-Q/cgn2.1/USPTO.spool_p/US10761006/runat_27122005_192900_15751/app_query.fasta_1.1102
-DB=Uniprot -QPMT=fastan -SUPFIX=rup -MINMATCH=0.1 -LOOPT=0 -LOOPT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10761006 @CGN 1.1.580 @runat_27122005_192900_15751 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Uniprot_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	139	100.0	226	2	Q98H30_HPBVO
2	139	100.0	226	2	Q98K9_HPBVO
3	139	100.0	400	2	Q80S15_HPBVO
4	136	97.8	71	2	Q76N5_HPBVO
5	136	97.8	71	2	Q9YXC2_HPBVO
6	136	97.8	98	2	Q9QKT6_HPBVO
7	136	97.8	215	2	Q9Q3D0_HPBVO
8	136	97.8	226	2	Q8BCB1_HPBVO
9	136	97.8	226	2	Q9H51_HPBVO
10	136	97.8	281	2	Q8BCB2_HPBVO
11	136	97.8	400	2	Q8BCB3_HPBVO
12	135	97.1	63	2	Q9Q517_HPBVO
13	135	97.1	226	2	Q80B7_HPBVO
14	135	97.1	226	2	Q9DUJ4_HPBVO
15	135	97.1	226	2	Q9DUK6_HPBVO
16	135	97.1	226	2	Q98LJ3_HPBVO

17	135	97.1	226	2	Q988L7_HPBVO	Q9817 hepatitis b
18	135	97.1	400	2	Q39645_HPBVO	Q3945 hepatitis b
19	135	97.1	400	2	Q91525_HPBVO	Q9125 hepatitis b
20	135	97.1	400	2	Q4FDM1_HPBVO	Q4fmi hepatitis b
21	134	96.4	47	2	Q9JFX8_HPBVO	Q9jfx8 hepatitis b
22	133	95.7	47	2	Q9JFX9_HPBVO	Q9jfx9 hepatitis b
23	133	95.7	53	2	Q71983_HPBVO	Q71984 hepatitis b
24	133	95.7	53	2	Q71984_HPBVO	Q71984 hepatitis b
25	133	95.7	53	2	Q71987_HPBVO	Q71987 hepatitis b
26	133	95.7	101	2	Q52X2M_HPBVO	Q52xm2 hepatitis b
27	133	95.7	101	2	Q52Y84_HPBVO	Q52ye4 hepatitis b
28	133	95.7	128	2	Q8QVC3_HPBVO	Q8qvc3 hepatitis b
29	133	95.7	128	2	Q8QVC4_HPBVO	Q8qvc4 hepatitis b
30	133	95.7	226	2	Q04258_HPBVO	Q04258 hepatitis b
31	133	95.7	226	2	Q91938_HPBVO	Q91938 hepatitis b
32	133	95.7	226	2	Q598Q0_HPBVO	Q598q0 hepatitis b
33	133	95.7	226	2	Q76U22_HPBVO	Q76uu2 duck hepati
34	133	95.7	226	2	Q9Q3C8_HPBVO	Q9q3c8 hepatitis b
35	133	95.7	226	2	Q9Q3E8_HPBVO	Q9q3e8 hepatitis b
36	133	95.7	226	2	Q9YD6_HPBVO	Q9yde6 hepatitis b
37	133	95.7	384	2	Q86986_HPBVO	Q86986 hepatitis b
38	133	95.7	389	2	Q91C55_HPBVO	Q91c55 hepatitis b
39	133	95.7	389	2	Q9Y2T8_HPBVO	Q9yzt8 hepatitis b
40	133	95.7	400	2	Q56U13_HPBVO	Q56u13 hepatitis b
41	133	95.7	400	2	Q5Y2B8_HPBVO	Q5y2b8 hepatitis b
42	130	93.5	42	2	Q8Q0M2_HPBVO	Q80qm2 hepatitis b
43	130	93.5	42	2	Q8Q0M4_HPBVO	Q80qm4 hepatitis b
44	130	93.5	42	2	Q8Q0M8_HPBVO	Q80qm8 hepatitis b
45	130	93.5	42	2	Q8Q0N0_HPBVO	Q80qn0 hepatitis b

ALIGNMENTS

RESULT 1

Q80H30_HPBVO
ID Q80H30_HPBVO PRELIMINARY; PRT; 226 AA.
AC Q80H30;
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE S protein.
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthonepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FMU005;
RA Lin X., Zheng D.L., Xu X.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY206376; AAP06547.1; -; Genomic DNA.
DR GO; GO:0016032; P:Viral life cycle; IEA.
DR InterPro; IPR000349; Hepvir_surfac.
DR PANTHER; PTHR10832; Hepvir_surfac; 1.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.
SQ SEQUENCE 226 AA; 25405 MW; 186C43D50FB8DAA CRC64;

Alignment Scores:
Pred. No.: 1,29e-13 Length: 226
Score: 139.00 Matches: 23
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x Q80H30_HPBVO (1-226)

QY 1 ACAGCTCTGCTCAGGAAACTCTAGCTTTCCCTCTGTGCTGTCACAAACCTTCGGAC 60
|||||
Db 125 ThrThrProAlaGlnGlyAanSerThrPheProSerCysCysCysThrLysProSerAsp 144
QY 61 GGAACCTGC 69

DR Pfam; PF00695; vNSA; 1.
 KW Antigen.
 SQ SEQUENCE 400 AA; 43741 MW; D1DFC54CDF11PA91 CRC64;

Alignment Scores:
 Pred. No.: 1.34e-13 Length: 400
 Score: 139.00 Matches: 23
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x Q80515_HPBVO (1-400)

Qy 1 ACGACTCTGCTCAAGAAACTCTAGCTTTCCCTCTTGTGTGTACAAAACCTTCGGAC 60
 |||||
 Db 299 ThrThrProAlaGlnGlyAsnSerThrPheProSerCysCysThrLysProSerAsp 318
 |||||

Qy 61 GGAACACTGC 69
 |||||
 Db 319 GlyAsnCys 321
 |||||

RESULT 4
 Q767N5_HPBVO
 ID Q767N5_HPBVO PRELIMINARY; PRT; 71 AA.
 AC Q767N5;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Surface antigen (Fragment).
 OS Hepatitis B virus.
 OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
 OC Orthohepadnavirus.
 OX NCBI_TaxID=10407;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Lusida M.I., Surayah, Sakugawa H., Nagano-Fujii M., Soetjipto,
 RA Mulyanto, Handayani R., Boediwarsono, Setiawan P.B., Nidom C.A.,
 RA Ohgimoto S., Hotta H.;
 RT "Genotype and Subtype Analyses of Hepatitis B Virus (HBV) and Possible
 RT Co-infection of HBV and Hepatitis C Virus (HCV) or Hepatitis D Virus
 RT (HDV) in Blood Donors, Patients with Chronic Liver Disease and
 RT Patients on Hemodialysis in Surabaya, Indonesia.";
 RL Microbiol. Immunol. 47:969-975 (2003).
 DR EMBL; AB113296; BAC99010.1; -; Genomic DNA.
 DR GO; GO:0016032; P: viral life cycle; IEA.
 DR InterPro; IPR000349; Hepvir_surfaG.
 DR PANTHER; PTHR10832; Hepvir_surfaG; 1.
 DR Pfam; PF00695; vNSA; 1.
 KW Antigen.
 FT NON_TER 1
 FT NON_TER 71
 SQ SEQUENCE 71 AA; 7861 MW; A8B0FB021386382C CRC64;

Alignment Scores:
 Pred. No.: 3.75e-13 Length: 71
 Score: 136.00 Matches: 22
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 95.65% Mismatches: 0
 Query Match: 97.84% Indels: 0
 DB: 2 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x Q767N5_HPBVO (1-71)

Qy 1 ACGACTCTGCTCAAGAAACTCTAGCTTTCCCTCTTGTGTGTACAAAACCTTCGGAC 60
 |||||
 Db 13 ThrThrProAlaGlnGlyAsnSerThrPheProSerCysCysThrLysProThrAsp 32
 |||||

Qy 61 GGAACACTGC 69
 |||||
 Db 33 GlyAsnCys 35
 |||||

RESULT 5

```

Q9YXC2_HPBVO
ID Q9YXC2_HPBVO PRELIMINARY; PRT; 71 AA.
AC Q9YXC2;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Surface antigen (Fragment).
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1] NUCLEOTIDE SEQUENCE.
RP MEDLINE=98449868; PubMed=9774595;
RX Mbayed V.A., Lopez J.L., Telenta P.F.S., Palacios G., Badia I.,
RA Ferro A., Galoppo C., Campos R.;
RT "Distribution of hepatitis B virus genotypes in two different
RL J. Clin. Microbiol. 36:3362-3365(1998).
RL EMBL; AF043576; AAC79977.1; -; Genomic DNA.
DR GO; GO:0016032; P:Viral life cycle; IEA.
DR InterPro; IPR000349; Hepvir_surfaG.
DR PANTHER; PTHR10832; Hepvir_surfaG; 1.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.
FT NON_TER 1 1
FT NON_TER 71 71
SQ SEQUENCE 71 AA; 7847 MW; B94A1002139B353C CRC64;

Alignment Scores:
Pred. No.: 3,75e-13 Length: 71
Score: 136.00 Matches: 22
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 95.65% Mismatches: 0
Query Match: 97.84% Indels: 0
DB: 2 Gaps: 0

US-10-761-006a-1_COPY_527_595 (1-69) x Q9YXC2_HPBVO (1-71)

QY 1 ACAGCTCTCTCAAGAACTCTAGTTTCCCTTGTGCTGACAAACCTTCGGAC 60
Db 13 ThrThrProAlaGlnGlyAenSerThrPheProSerCysCysThrLysProThrAsp 32

QY 61 GGAACCTGC 69
Db 33 GlyAenCys 35

RESULT 6
Q9QKT6_HPBVO
ID Q9QKT6_HPBVO PRELIMINARY; PRT; 98 AA.
AC Q9QKT6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE HBsAg (Fragment).
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1] NUCLEOTIDE SEQUENCE.
RP Swenson P.D., Van Geyt C., Alexander R.E., Hagan H.,
RA Freitag-Koontz J.M., De Gendt S., Van Reybroeck G., Wilson S.,
RA Magnus L., Stuyver L.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF107159; AAF04986.1; -; Genomic DNA.
DR GO; GO:0016032; P:Viral life cycle; IEA.
DR InterPro; IPR000349; Hepvir_surfaG.
DR PANTHER; PTHR10832; Hepvir_surfaG; 1.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.
FT NON_TER 1 1
FT NON_TER 98 98
SQ SEQUENCE 98 AA; 11040 MW; 09B458DE7C0BB075 CRC64;

Alignment Scores:
Pred. No.: 4,04e-13 Length: 215
Score: 136.00 Matches: 22
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 95.65% Mismatches: 0
Query Match: 97.84% Indels: 0
DB: 2 Gaps: 0

US-10-761-006a-1_COPY_527_595 (1-69) x Q9QKT6_HPBVO (1-215)

QY 1 ACAGCTCTCTCAAGAACTCTAGTTTCCCTTGTGCTGACAAACCTTCGGAC 60
Db 125 ThrThrProAlaGlnGlyAenSerThrPheProSerCysCysThrLysProThrAsp 144

QY 61 GGAACCTGC 69
Db 145 GlyAenCys 147

RESULT 8
Q8BCB1_HPBVO
ID Q8BCB1_HPBVO PRELIMINARY; PRT; 226 AA.
AC Q8BCB1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

```

```
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE S protein.
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22652772; PubMed=12767980;
RX DOI=10.1128/JVI.77.12.6601-6612.2003;
RA Parekh S., Zoulim F., Abn S.H., Tsai A., Li J., Kawai S., Khan N.,
RA Trepo C., Wands J., Tong S.;
RT "Genome replication, virion secretion, and e antigen expression of
RT naturally occurring hepatitis B virus core promoter mutants.";
RL J. Virol. 77:6601-6612(2003).
DR EMBL; AF537372; AN15124.1; -; Genomic DNA.
DR GO; GO:0016032; P:Viral life cycle; IEA.
DR InterPro; IPR000349; Hepvir_surfa; IEA.
DR PANTHER; PTHR10832; Hepvir_surfa; 1.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.
SQ SEQUENCE 226 AA; 25433 MW; CBCFB5CAFAA82107 CRC64;

Alignment Scores:
Pred. No.: 4.05e-13 Length: 226
Score: 136.00 Matches: 22
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 95.65% Mismatches: 0
Query Match: 97.84% Indels: 0
DB: 2 Gaps: 0

US-10-761-006a-1_COPY_527_595 (1-69) x Q8BCB1_HPBVO (1-226)

QY 1 AGCACTCTGCTCAAGAACTCTAGTTTCCCTCTGTGTGCTGTACAAACCTTCGGAC 60
|||||
Db 125 ThrThrProAlaGlnGlyAsnSerThrPheProSerCysCysCysThrLysProThrAsp 144

QY 61 GGAAACTGC 69
|||||
Db 145 GlyAsnCys 147

RESULT 9
Q8BCB2_HPBVO
ID Q8BCB2_HPBVO PRELIMINARY; PRT; 281 AA.
AC Q8BCB2;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Middle S protein.
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22652772; PubMed=12767980;
RX DOI=10.1128/JVI.77.12.6601-6612.2003;
RA Parekh S., Zoulim F., Abn S.H., Tsai A., Li J., Kawai S., Khan N.,
RA Trepo C., Wands J., Tong S.;
RT "Genome replication, virion secretion, and e antigen expression of
RT naturally occurring hepatitis B virus core promoter mutants.";
RL J. Virol. 77:6601-6612(2003).
DR EMBL; AF537372; AN15128.1; -; Genomic DNA.
DR GO; GO:0016032; P:Viral life cycle; IEA.
DR InterPro; IPR000349; Hepvir_surfa; 1.
DR PANTHER; PTHR10832; Hepvir_surfa; 1.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.
SQ SEQUENCE 281 AA; 31149 MW; 7CA77FAEB18CDBC9 CRC64;

Alignment Scores:
Pred. No.: 4.11e-13 Length: 281
Score: 136.00 Matches: 22
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 95.65% Mismatches: 0
Query Match: 97.84% Indels: 0
DB: 2 Gaps: 0

US-10-761-006a-1_COPY_527_595 (1-69) x Q8BCB2_HPBVO (1-281)

QY 1 AGCACTCTGCTCAAGAACTCTAGTTTCCCTCTGTGTGCTGTACAAACCTTCGGAC 60
|||||
Db 180 ThrThrProAlaGlnGlyAsnSerThrPheProSerCysCysCysThrLysProThrAsp 199

QY 61 GGAAACTGC 69
```

```
Db      200 GlyAsnCys 202
|||||
RESULT 11
Q8BCB3_HPBVO
ID Q8BCB3_HPBVO PRELIMINARY; PRT; 400 AA.
AC Q8BCB3
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Large S protein.
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22652772; PubMed=12767980;
RX DOI=10.1128/JVI.77.12.6601-6612.2003;
RA Parekh S., Zoulim F., Ahn S.H., Tsai A., Li J., Kawai S., Khan N.,
RA Trepo C., Wands J., Tong S.;
RT "Genome replication, virion secretion, and e antigen expression of
RT naturally occurring hepatitis B virus core promoter mutants.";
RL J. Virol. 77:6601-6612(2003).
DR EMBL; AF537372; AANL5127.1; -; Genomic_DNA.
DR GO; GO:0016032; P:Viral life cycle; IEA.
DR InterPro; IPR000349; Hepvir_surfaG.
DR PANTHER; PTHR10832; Hepvir_surfaG; 1.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.
SQ SEQUENCE 400 AA; 43651 MW; 6E6C6630699CC81E CRC64;

Alignment Scores:
Pred. No.: 4.21e-13 Length: 400
Score: 136.00 Matches: 22
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 95.65% Mismatches: 0
Query Match: 97.84% Indels: 0
DB: 2 Gaps: 0

US-10-761-006a-1_COPY_527_595 (1-69) x Q8BCB3_HPBVO (1-400)

Qy 1 ACAGCTCCTGCTCAAGAACTCTAGCTTCCCTCTGTGCTGTACAAACCTTCGGAC 60
Db 299 ThrThrProAlaGlnGlyAsnSerThrPheProSerCysCysThrLysProThrAsp 318
Qy 61 GGAAACTGC 69
Db 319 GlyAsnCys 321

RESULT 12
Q9Q5L7_HPBVO
ID Q9Q5L7_HPBVO PRELIMINARY; PRT; 63 AA.
AC Q9Q5L7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Surface antigen (Fragment).
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Xia G.-L., Nainan O.V., Jia Z.-Y., Cao H.-L., Liu C.-B.,
RA Margolis H.S.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF198960; AAF15879.1; -; Genomic DNA.
DR GO; GO:0016032; P:Viral life cycle; IEA.
DR InterPro; IPR000349; Hepvir_surfaG.
DR PANTHER; PTHR10832; Hepvir_surfaG; 1.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.

Alignment Scores:
Pred. No.: 4.21e-13 Length: 400
Score: 136.00 Matches: 22
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 95.65% Mismatches: 0
Query Match: 97.84% Indels: 0
DB: 2 Gaps: 0

US-10-761-006a-1_COPY_527_595 (1-69) x Q8BCB3_HPBVO (1-400)

Qy 1 ACAGCTCCTGCTCAAGAACTCTAGCTTCCCTCTGTGCTGTACAAACCTTCGGAC 60
Db 299 ThrThrProAlaGlnGlyAsnSerThrPheProSerCysCysThrLysProThrAsp 318
Qy 61 GGAACTGC 69
Db 319 GlyAsnCys 321

RESULT 13
Q80BT7_HPBVO
ID Q80BT7_HPBVO PRELIMINARY; PRT; 226 AA.
AC Q80BT7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE S protein.
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=95154763; PubMed=7851832;
RX Moriyama K., Takada T., Tsutsumi Y., Fukada K., Ishibashi H., Niho Y.,
RX Maeda Y.;
RT "Mutations in the transcriptional regulatory region of the precore and
RT core/pregenome of a hepatitis B virus with defective HBeAg
RT production.";
RL Fukuoka Igaku Zasshi 85:314-322(1994).
DR EMBL; S75184; AAP31568.1; -; Genomic_DNA.
DR GO; GO:0016032; P:Viral life cycle; IEA.
DR InterPro; IPR000349; Hepvir_surfaG.
DR PANTHER; PTHR10832; Hepvir_surfaG; 1.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.
SQ SEQUENCE 226 AA; 25365 MW; BC8PE3999F34789F CRC64;

Alignment Scores:
Pred. No.: 5.93e-13 Length: 226
Score: 135.00 Matches: 22
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 95.65% Mismatches: 0
Query Match: 97.12% Indels: 0
DB: 2 Gaps: 0

US-10-761-006a-1_COPY_527_595 (1-69) x Q80BT7_HPBVO (1-226)

Qy 1 ACAGCTCCTGCTCAAGAACTCTAGCTTCCCTCTGTGCTGTACAAACCTTCGGAC 60
Db 125 ThrSerProAlaGlnGlyAsnSerThrPheProSerCysCysThrLysProSerAsp 144
Qy 61 GGAACTGC 69
Db 145 GlyAsnCys 147

RESULT 14
Q9DUJ4_HPBVO
ID Q9DUJ4_HPBVO PRELIMINARY; PRT; 226 AA.
AC Q9DUJ4;
```

```
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Small surface polypeptide.
GN Name=OG14;
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Ogura Y., Kurosaki M., Asahina Y., Enomoto N., Marumo F., Sato C.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB029963; BAB20346.1; -; Genomic_DNA.
DR PIR; JQ2094; JQ2094.
DR PIR; JQ2095; JQ2095.
DR PIR; JQ2096; JQ2096.
DR PIR; JQ2097; JQ2097.
DR PIR; JQ2098; JQ2098.
DR PIR; JQ2099; JQ2099.
DR PIR; JQ2100; JQ2100.
DR PIR; JQ2101; JQ2101.
DR PIR; JQ2102; JQ2102.
DR PIR; JQ2106; JQ2106.
DR PIR; JQ2108; JQ2108.
DR PIR; JQ2109; JQ2109.
DR PIR; JQ2111; JQ2111.
DR PIR; JQ2112; JQ2112.
DR PIR; JQ2116; JQ2116.
DR GO; GO:0016032; P:Viral life cycle; IEA.
DR InterPro; IPR000349; Hepvir_surfaG.
DR PANTHER; PTHR10832; Hepvir_surfaG; 1.
DR Pfam; PF00695; vMSA; 1.
KW Antigen; Polypeptide.
SQ SEQUENCE 226 AA; 25375 MW; 7704565C4092E3BC CRC64;

Alignment Scores:
Pred. No.: 5.93e-13 Length: 226
Score: 135.00 Matches: 22
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 95.65% Mismatches: 0
Query Match: 97.12% Indels: 0
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x Q9DUJ4_HPBVO (1-226)

QY 1 ACGACTCTGCTCAAGGAACCTACGTTCCCTCTTGTGTACAAAACCTTCGGAC 60
DB 125 ThrSerProAlaGlnGlyAsnSerThrPheProSerCysCysThrLysProSerAsp 144

QY 61 GGAAACTGC 69
DB 145 GlyAsnCys 147

RESULT 15
Q9DUK6_HPBVO
ID Q9DUK6_HPBVO PRELIMINARY; PRT; 226 AA.
AC Q9DUK6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Small surface polypeptide.
GN Name=OG2;
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Ogura Y., Kurosaki M., Asahina Y., Enomoto N., Marumo F., Sato C.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB029951; BAB20334.1; -; Genomic_DNA.
DR PIR; JQ2102; JQ2102.

Alignment Scores:
Pred. No.: 5.93e-13 Length: 226
Score: 135.00 Matches: 22
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 95.65% Mismatches: 0
Query Match: 97.12% Indels: 0
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x Q9DUJ4_HPBVO (1-226)

QY 1 ACGACTCTGCTCAAGGAACCTACGTTCCCTCTTGTGTACAAAACCTTCGGAC 60
DB 125 ThrSerProAlaGlnGlyAsnSerThrPheProSerCysCysThrLysProSerAsp 144

QY 61 GGAAACTGC 69
DB 145 GlyAsnCys 147

Search completed: December 27, 2005, 21:00:25
Job time : 24.046 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 27, 2005, 20:44:58 ; Search time 45.4 Seconds
(without alignments)
2480.270 Million cell updates/sec

Title: US-10-761-006A-1_COPY_155_835
Perfect score: 1278
Sequence: 1 ATGGAGACACACATCAGG.....GCTTTGGGTATACATTAA 681

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US10761006/runat_27122005.192901.15775/app_query.fasta_1.1102
-DB=Issued Patents AA -OPMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=spt -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10761006 @CIGN 1 1 01 @runat_27122005.192901.15775 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5 COMB.pcp:*
2: /cgn2_6/ptodata/1/iaa/6 COMB.pcp:*
3: /cgn2_6/ptodata/1/iaa/H COMB.pcp:*
4: /cgn2_6/ptodata/1/iaa/PCUS COMB.pcp:*
5: /cgn2_6/ptodata/1/iaa/RE COMB.pcp:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pcp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1230	96.2	400	2	US-09-719-528A-3
2	1230	96.2	400	2	US-10-209-264-3
3	1206	94.4	226	1	US-08-378-011A-3
4	1206	94.4	236	1	US-08-378-011A-1
5	1199	93.8	277	6	5164485-2
6	1190	93.1	843	2	US-09-719-528A-2
7	1190	93.1	843	2	US-10-209-264-2
8	1189	93.0	226	2	US-09-471-573A-2
9	1189	93.0	226	2	US-09-471-573A-40
10	1163	91.0	226	6	5196194-21
11	1163	91.0	226	6	5436139-5
12	1163	91.0	390	2	US-09-721-480-5

13	1163	91.0	531	2	US-09-721-480-7
14	1160	90.8	423	1	US-08-760-797A-1
15	1160	90.8	424	1	US-08-760-797A-3
16	1160	90.8	424	2	US-08-932-929B-1
17	1160	90.8	424	2	US-08-932-929B-3
18	1157	90.5	281	2	US-09-247-890-10
19	1157	90.5	281	2	US-09-724-969-10
20	1157	90.5	281	2	US-09-724-852-10
21	1156	90.5	281	2	US-09-721-480-3
22	1152	90.1	226	6	5198348-1
23	1151	90.1	225	6	5436139-4
24	1143	89.5	395	6	5196194-18
25	1133	88.7	226	2	US-10-104-966-10
26	1133	88.7	226	2	US-09-929-955-10
27	1126	88.1	226	4	PCT-US96-10602-14
28	1126	88.1	281	1	US-08-105-483-214
29	1126	88.1	281	1	US-08-709-209-214
30	1126	88.1	281	1	US-08-458-101-214
31	1126	88.1	281	2	US-09-247-890-12
32	1126	88.1	281	2	US-09-724-969-12
33	1126	88.1	281	2	US-09-724-852-12
34	1126	88.1	389	1	US-08-105-483-216
35	1126	88.1	389	1	US-08-105-483-219
36	1126	88.1	389	1	US-08-709-209-216
37	1126	88.1	389	1	US-08-709-209-219
38	1126	88.1	389	1	US-08-458-101-216
39	1126	88.1	389	1	US-08-458-101-219
40	1126	88.1	389	2	US-08-486-099-106
41	1126	88.1	389	2	US-08-360-107A-116
42	1126	88.1	389	2	US-08-484-223B-106
43	1126	88.1	389	2	US-08-919-597-106
44	1126	88.1	389	2	US-08-475-668A-106
45	1126	88.1	389	2	US-08-485-551A-106

ALIGNMENTS

RESULT 1
US-09-719-528A-3
; Sequence 3, Application US/09719528A
; Patent No. 6558675

; GENERAL INFORMATION:

; APPLICANT: Oon, Chong Jin
; Lim, Gek Keow

; Zhao, Yi

; Chen, Wei Ning

; TITLE OF INVENTION: A MUTANT HUMAN HEPATITIS B VIRAL STRAIN AND

; USES THEREOF

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESSES:

; ADDRESSEE: Ladas & Parry

; STREET: 26 West 61 Street

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10023

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/719,528A

; FILING DATE: 30-Apr-2001

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/SG98/00046

; FILING DATE: 19-JAN-1998

; ATTORNEY/AGENT INFORMATION:

; NAME: Mass, Clifford J.

; REGISTRATION NUMBER: 30,086

; REFERENCE/DOCKET NUMBER: U-013109-7

; TELECOMMUNICATION INFORMATION:

Sequence 7, Appli
Sequence 1, Appli
Sequence 3, Appli
Sequence 1, Appli
Sequence 3, Appli
Sequence 10, Appli
Sequence 10, Appli
Sequence 10, Appli
Sequence 3, Appli
Patent No. 5198348
Patent No. 5436139
Patent No. 5196194
Sequence 10, Appli
Sequence 10, Appli
Sequence 14, Appli
Sequence 214, App
Sequence 214, App
Sequence 214, App
Sequence 12, Appli
Sequence 12, Appli
Sequence 12, Appli
Sequence 216, App
Sequence 219, App
Sequence 216, App
Sequence 219, App
Sequence 216, App
Sequence 106, App
Sequence 116, App
Sequence 106, App
Sequence 106, App
Sequence 106, App

```

;
; TELEPHONE: (212) 708-1800
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 400 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-719-528A-3

Alignment Scores:
Pred. No.: 3.74e-124 Length: 400
Score: 1230.00 Matches: 224
Percent Similarity: 98.68% Conservative: 0
Best Local Similarity: 98.68% Mismatches: 2
Query Match: 96.24% Indels: 2
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x US-09-719-528A-3 (1-400)
QY 1 ATGGAGAACACACATCAGGATTCCTAGGACCCCTGCTGCTTACAGCGGGGTTTTTC 60
Db 175 MetGluAsnThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 194
QY 61 TCGTTGACAGAATCCTCAATACCGCAGAGTCTAGACTC-TGGTGGAGTCTCTCTCAAT 119
Db 195 SerLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 214
QY 120 TTTCTAGGGGAGCAGCCACGCTGTTCTGGCCAAATTCGAGTCCCACTCCCAATCA 179
Db 215 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThr-SerAsnHi 234
QY 180 CTCACCAACCTCTTGTCTCCCAATTGCTCGCTATCGCTGATGTGTCTGGGGGTTT 239
Db 234 sSerProThrSerCysProProlleCysProGlyTyrArgTrpAsnCysLeuArgGph 254
QY 240 TATCATATTCTCTTCATCTGCTGCTATGCTCATCTTCTTGTGTTCTCTGACTA 299
Db 254 ellellePheLeuPheilleuLeuLeuLeuLeuLeuLeuValLeuLeuAspTy 274
QY 300 CCAAGGTATGTTGCCCGTTTGTCTTCTACTTCCAGGAACATCAACACGACGCGGCC 359

US-10-209-264-3
Alignment Scores:
Pred. No.: 3.74e-124 Length: 400
Score: 1230.00 Matches: 224
Percent Similarity: 98.68% Conservative: 0
Best Local Similarity: 98.68% Mismatches: 2
Query Match: 96.24% Indels: 2
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x US-10-209-264-3 (1-400)
QY 1 ATGGAGAACACACATCAGGATTCCTAGGACCCCTGCTGCTTACAGCGGGGTTTTTC 60
Db 175 MetGluAsnThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 194
QY 61 TCGTTGACAGAATCCTCAATACCGCAGAGTCTAGACTC-TGGTGGAGTCTCTCTCAAT 119
Db 195 SerLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 214
QY 120 TTTCTAGGGGAGCAGCCACGCTGTTCTGGCCAAATTCGAGTCCCACTCCCAATCA 179
Db 215 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThr-SerAsnHi 234
QY 180 CTCACCAACCTCTTGTCTCCCAATTGCTCGCTATCGCTGATGTGTCTGGGGGTTT 239
Db 234 sSerProThrSerCysProProlleCysProGlyTyrArgTrpAsnCysLeuArgGph 254
QY 240 TATCATATTCTCTTCATCTGCTGCTATGCTCATCTTCTTGTGTTCTCTGACTA 299
Db 254 ellellePheLeuPheilleuLeuLeuLeuLeuLeuLeuValLeuLeuAspTy 274
QY 300 CCAAGGTATGTTGCCCGTTTGTCTTCTACTTCCAGGAACATCAACACGACGCGGCC 359

US-10-209-264-3
Alignment Scores:
Pred. No.: 3.74e-124 Length: 400
Score: 1230.00 Matches: 224
Percent Similarity: 98.68% Conservative: 0
Best Local Similarity: 98.68% Mismatches: 2
Query Match: 96.24% Indels: 2
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x US-10-209-264-3 (1-400)
QY 1 ATGGAGAACACACATCAGGATTCCTAGGACCCCTGCTGCTTACAGCGGGGTTTTTC 60
Db 175 MetGluAsnThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 194
QY 61 TCGTTGACAGAATCCTCAATACCGCAGAGTCTAGACTC-TGGTGGAGTCTCTCTCAAT 119
Db 195 SerLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 214
QY 120 TTTCTAGGGGAGCAGCCACGCTGTTCTGGCCAAATTCGAGTCCCACTCCCAATCA 179
Db 215 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThr-SerAsnHi 234
QY 180 CTCACCAACCTCTTGTCTCCCAATTGCTCGCTATCGCTGATGTGTCTGGGGGTTT 239
Db 234 sSerProThrSerCysProProlleCysProGlyTyrArgTrpAsnCysLeuArgGph 254
QY 240 TATCATATTCTCTTCATCTGCTGCTATGCTCATCTTCTTGTGTTCTCTGACTA 299
Db 254 ellellePheLeuPheilleuLeuLeuLeuLeuLeuLeuValLeuLeuAspTy 274
QY 300 CCAAGGTATGTTGCCCGTTTGTCTTCTACTTCCAGGAACATCAACACGACGCGGCC 359

```

RESULT 2

US-10-209-264-3
; Sequence 3, Application US/10209264
; Patent No. 6787142
; GENERAL INFORMATION:

Db 274 rGInGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrGlyPr 294
Qy 360 ATGCAAGACTGACGACTCTGCTCAAGAACTCTACGTTCCCTCTTGTGCTGTAC 419
Db 294 oCysLysThrCysThrThrProAlaGInGlyAsnSerThrPheProSerCysCysTh 314
Qy 420 AAAACCTTCGACGGAACGCACTCTATTCCCATCCCATCTCTGGGCTTTCGCAAG 479
Db 314 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerThrPheAla 334
Qy 480 ATTCCTATGGAGTGGCCCTCAGTCCGTTTCTCTCGCTCAGTTTACTAGTGCATTGT 539
Db 334 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValPro 354
Qy 540 TCAGTGTTCGTAGGCTTTCCTCCCACTGTTTGGCTTTCAGTTATATGATGATG 599
Db 354 lGInTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetTr 374
Qy 600 TTGGGGCGGAAGTCTGTACAACTCTTGAGTCCCTTTTACCTCTATTACCAATTTCTT 659
Db 374 rTrpGlyArgSerLeuTrpAsnIleLeuSerProPheLeuProLeuLeuProIle 394
Qy 660 TTGTCTTTGGGTATACATT 678
Db 394 eCysLeuTrpValTrpIle 400

RESULT 3

US-08-378-011A-3
; Sequence 3, Application US/08378011A
; Patent No. 5693497
; GENERAL INFORMATION:
; APPLICANT: TAKAMIZAWA, Akihisa
; APPLICANT: FUJITA, Hiroyuki
; APPLICANT: MANABE, Sadao
; APPLICANT: KATO, Masahiko
; APPLICANT: OSAME, Juichiro
; APPLICANT: YOSHIDA, Iwao
; APPLICANT: KONOBE, Takeo
; APPLICANT: TAKAKU, Keisuke
; TITLE OF INVENTION: HEPATITIS B VIRUS ANTIGEN AND A PROCESS FOR
; TITLE OF INVENTION: PRODUCING SAME
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII, generated using Word Perfect, version 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/378,011A
; FILING DATE: 25-JAN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 61-143412
; FILING DATE: 18-JUN-1986
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/061,518
; FILING DATE: 15-JUN-1987
; APPLICATION DATA:
; APPLICATION NUMBER: US 07/902,494
; FILING DATE: 23-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: 870602B
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 226 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-378-011A-3
Alignment Scores:
Pred. No.: 1,18e-121 Length: 226
Score: 1206.00 Matches: 220
Percent Similarity: 96.92% Conservative: 0
Best Local Similarity: 96.92% Mismatches: 6
Query Match: 94.37% Indels: 2
DB: 1 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x US-08-378-011A-3 (1-226)

Qy 1 ATGAGAACACAACTCAGGATTCCTAGGACCCCTCTCTGTGTACAGGCGGGTTTTTC 60
Db 1 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20
Qy 61 TCGTTGACAAGATCCTCACATACCGCAGACTAGACTG-TGGTGGACTTCTCTCAAT 119
Db 21 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 40
Qy 120 TTTCTAGGGGAGCACCACGCTGTTCTGGCCAAAATTCAGTCCCAACCTCCCAATCA 179
Db 41 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 60
Qy 180 CTCACCAACTCTTGTCTCTCAATTTGTCTGCTATCGCTGGATGTCGTGGCGCTTT 239
Db 60 sSerProThrSerCysProPheCysProGlyTyArgTrpMetCysLeuArgArgPh 80
Qy 240 TATCATATTCCTCTCTCATCTGCTATGCTCATCTCTTCTTGTGGTCTTCTGGACTA 299
Db 80 eIleIlePheLeuPheIleLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTy 100
Qy 300 CCAAGGTATGTTGCCGTTTGTCTCTTACTTCCAGGAACATCAACACAGCAGCGGGGCC 359
Db 100 rGInGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 120
Qy 360 ATGCAAGACTCAGCAGACTCTCTCTCAAGGAAACTCTAGCTTTCCCTCTTGTGCTGTAC 419
Db 120 oCysLysThrCysThrIleProAlaGInGlyThrSerMetPheProSerCysCysTh 140
Qy 420 AAAACCTTCGACGGAACGCACTTGTATTCATCCATCCCATCATCTGGGCTTTCGCAAG 479
Db 140 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerThrPheAla 160
Qy 480 ATTCCTATGGAGTGGCCCTCAGTCCGTTTCTCTCGCTCAGTTTACTAGTGCATTGT 539
Db 160 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPhe 180
Qy 540 TCAGTGTTCGTAGGCTTTCCTCCCACTGTTTGGCTTTCAGTTTATATGATGATG 599
Db 180 lGInTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetTr 200
Qy 600 TTGGGGCGGAAGTCTGTACAACTCTTGAGTCCCTTTTACCTCTATTACCAATTTCTT 659
Db 200 rTrpGlyProSerLeuTrpAsnIleLeuSerProPheLeuProLeuLeuProIlePhe 220
Qy 660 TTGTCTTTGGGTATACATT 678
Db 220 eCysLeuTrpValTrpIle 226

RESULT 4

US-08-378-011A-1
; Sequence 1, Application US/08378011A
; Patent No. 5693497

```

; GENERAL INFORMATION:
; APPLICANT: TAKAMIZAWA, Akihisa
; APPLICANT: FUJITA, Hiroyuki
; APPLICANT: MANABE, Sadao
; APPLICANT: KATO, Masahiko
; APPLICANT: OSAME, Juichiro
; APPLICANT: YOSHIDA, Iwao
; APPLICANT: KONOE, Takeo
; APPLICANT: TAKAKU, Keisuke
; TITLE OF INVENTION: HEPATITIS B VIRUS ANTIGEN AND A PROCESS FOR
; TITLE OF INVENTION: PRODUCING SAME
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McIeland &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII, generated using Word Perfect, version 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/378,011A
; FILING DATE: 25-JAN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 61-143412
; FILING DATE: 18-JUN-1986
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/061,518
; FILING DATE: 15-JUN-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/902,494
; FILING DATE: 23-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: 870602B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 236 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-378-011A-1

Alignment Scores:
Pred. No.: 1,2e-121 Length: 236
Score: 1206.00 Matches: 220
Percent Similarity: 96.92% Conservative: 0
Best Local Similarity: 96.92% Mismatches: 6
Query Match: 94.37% Indels: 2
DB: 1 Gaps: 0

US-10-761-006a-1_COPY_155_835 (1-681) x US-08-378-011A-1 (1-236)

QY 1 ATGGAGAACACACATCAGATTCCTAGGACCCCTGCTGTTACAGCGGGGTTTTTC 60
Db 11 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 30
QY 61 TCGTTACACAGATCCTCAATACCGCAGAGTCTAGACTC-TGGTGGACTTCTCTCAAT 119
Db 31 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 50
QY 120 TTTCTAGGGGGAGCACCCACCGTGTTCCTCGCCAAAATTCCGAGTCCCCCAACCTCCAATCA 179

```

```

Db 51 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 70
QY 180 CTCACCAACCTCTGTCTCCTCAATTTGTCTCGGCTATCGCTGGATGTCTCTCGCGGCTTT 239
Db 70 sSerProThrSerCysProIleCysProGlyTrpArgTrpMetCysLeuArgArgPh 90
QY 240 TATCATATTCCTCTTTCATCTCTGCTATGCTCATCTTCTTGTGGTCTTCTTGACTA 299
Db 90 eIleIlePheLeuPheIleLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTy 110
QY 300 CCAAGGTATGTTGCCGTTTGTCTCTACTTCTCAGGAAACATCAACACACAGCGGGGCC 359
Db 110 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 130
QY 360 ATGCAAGACCTGCGACGACTCTCTGCTCAAGGAAACCTCTACGTTTCCCTCTTGTGCTG 419
Db 130 oCysLysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysTh 150
QY 420 AAAACCTTCGGACGGAAACTGCACCTTGTATTCCCATCCCATCATCTCTGGGCTTTCGCAAG 479
Db 150 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaAr 170
QY 480 ATTCTATGCGAGTGGGCGCTCAGTCCGTTTCTCTCGGCTCAGTTTACTAGTCCCATTTGT 539
Db 170 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa 190
QY 540 TCAGTGGTTCGTAGGGCTTCCCCACACTGTTGGCTTTCAGTTATATGATGATGTGTA 599
Db 190 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 210
QY 600 TTGGGGCGGAAAGTCTGTACAACATCTTGAGTCCCTTTTACCTCTATTATCCAAATTTCTT 659
Db 210 rTrpGlyProSerLeuTyAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh 230
QY 660 TTGTCTTTGGGTATACATT 678
Db 230 eCysLeuTrpValTyIle 236

RESULT 5
5164485-2
; Patent No. 5164485
; APPLICANT: FUJISAWA, YUKIO; ITOH, YASUAKI; NISHIMURA, OSAMU
; FUJII, TOMOKO
; TITLE OF INVENTION: MODIFIED HEPATITIS B VIRUS SURFACE
; ANTIGEN P31 AND PRODUCTION THEREOF
; NUMBER OF SEQUENCES: 22
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/547,948
; FILING DATE: 03-JUL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 898,425
; FILING DATE: 20-AUG-1986
; SEQ ID NO:2:
; LENGTH: 277
; 5164485-2

Alignment Scores:
Pred. No.: 7,27e-121 Length: 277
Score: 1199.00 Matches: 220
Percent Similarity: 96.92% Conservative: 0
Best Local Similarity: 96.92% Mismatches: 6
Query Match: 93.82% Indels: 2
DB: 1 Gaps: 0

US-10-761-006a-1_COPY_155_835 (1-681) x 5164485-2 (1-277)

QY 1 ATGGAGAACACACATCAGATTCCTAGGACCCCTGCTGTTACAGCGGGGTTTTTC 60
Db 52 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 71
QY 61 TCGTTGACACAGATCCTCAATACCGCAGAGTCTAGACTC-TGGTGGACTTCTCTCAAT 119

```



```
RESULT 7
US-10-209-264-2
; Sequence 2, Application US/10209264
; Patent No. 6787142
; GENERAL INFORMATION:
; APPLICANT: Oon, Chong Jin
; Lim, Gek Keow
; Zhao, Yi
; Chen, Wei Ning
; TITLE OF INVENTION: A MUTANT HUMAN HEPATITIS B VIRAL STRAIN AND
; USES THEREOF
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ladas & Parry
; STREET: 26 West 61 Street
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10023
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/209,264
; FILING DATE: 31-Jul-2002
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/SG98/00046
; FILING DATE: 19-JAN-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Mass, Clifford J.
; REGISTRATION NUMBER: 30,086
; REFERENCE/DOCKET NUMBER: U-013109-7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 708-1800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 843 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-209-264-2
Alignment Scores:
Pred. No.: 1,04e-119 Length: 843
Score: 1190.00 Matches: 226
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 93.11% Indels: 0
DB: 2 Gaps: 0
US-10-761-006a-1_COPY_155_835 (1-681) x US-10-209-264-2 (1-843)
QY 3 GGAGAACACATCAGGATTCCTAGGACCCCTCGTGTGTACAGCGGGGTTTTCTC 62
Db 356 GlyGluHisAsnIleArgIleProArgThrProAlaArgValThrGlyGlyValPheLeu 375
QY 63 GTTGACAGATCCTCACATACCGCAGAGTCTAGACTCTGGTGGACTTCTCTCAATTTT 122
Db 376 ValAspLysAsnProHisAsnThrAlaGluSerArgLeuTrpTrpThrSerLeuAsnPhe 395
QY 123 CTAGGGGAGCACCACGCTGCTCTGCGCAAAATTCGCAGTCCCAACCTCCAATCACATC 182
Db 396 LeuGlyGlyAlaProThrCysSerTrpProLysPheAlaValProAsnLeuGlnSerLeu 415
QY 183 ACCAACCTCTGTCTCCCAATTTGTCCTGGCTATCGCTGGATGTGCTCGCGGGTTTTAT 242
Db 416 ThrAsnLeuLeuSerSerAsnLeuSerTrpLeuSerLeuAspValSerAlaAlaPheTyr 435
QY 243 CATATTCCCTTCATCTCTGCTGCTATGCCATCTCTTCTTGTGGTCTTCTGCACTACCA 302
```

```
Db 436 HisIleProLeuHisProAlaAlaMetProHisLeuLeuValGlySerSerGlyLeuPro 455
QY 303 AGGTATGTTGCCCGTTTCTCTCTACTTCCAGGAACATCAACACACAGCAGCGGGCCATG 362
Db 456 ArgTyrValAlaArgLeuSerSerThrSerArgAsnIleAsnHisGlnHisGlyAlaMet 475
QY 363 CAAGACCTGCAGACTCCTGCTCAAGAAACTCTACGTTTCCCTCTTGTGCTGTACAAA 422
Db 476 GlnAspLeuHisAspSerCysSerArgLysLeuTyrValSerLeuLeuLeuTyrLys 495
QY 423 ACCTTCGACGAGAAACTGCACCTTGATTCCCATCCCATCATCTCGGCTTTTCGCAAGATT 482
Db 496 ThrPheGlyArgLysLeuHisLeuTyrSerHisProIleLeuGlyPheArgLysIle 515
QY 483 CCTATGGAGTGGGCTCAGTCCGCTTCTCCTGGCTCAGTTTACTAGTCCCATTTGTTC 542
Db 516 ProMetGlyValGlyLeuSerProPheLeuLeuAlaGlnPheThrSerAlaIleCysSer 535
QY 543 GTGGTTCGTAGGGCTTTCCCCACATGTTTGGCTTTTCACTATATGATGATGATGATTG 602
Db 536 ValValArgArgAlaPheProHisCysLeuAlaPheSerTyrMetAspValValLeu 555
QY 603 GGGCGAAGTCTGTACAACTCTGAGTCCCTTTTACTCTCTATTACCAATTTCTTTTG 662
Db 556 GlyAlaLysSerValGlnHisLeuGluSerLeuPheThrSerIleThrAsnPheLeu 575
QY 663 TCTTTGGGTATACATTTA 680
Db 576 SerLeuGlyIleHisLeu 581
RESULT 8
US-09-471-573A-2
; Sequence 2, Application US/09471573A
; Patent No. 6551820
; GENERAL INFORMATION:
; APPLICANT: Mason, Hugh
; APPLICANT: Thanavala, Yasmin
; TITLE OF INVENTION: Expression of Immunogenic Hepatitis B Surface Antigens In Transge
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 3121/1080
; CURRENT APPLICATION NUMBER: US/09/471,573A
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 60/113,827
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Wild-type HBeAg amino acid sequence
; NAME/KEY: misc_feature
; OTHER INFORMATION: Wild-type HBeAg amino acid sequence
US-09-471-573A-2
Alignment Scores:
Pred. No.: 8,1e-120 Length: 226
Score: 1189.00 Matches: 218
Percent Similarity: 96.04% Conservative: 0
Best Local Similarity: 96.04% Mismatches: 8
Query Match: 93.04% Indels: 2
DB: 2 Gaps: 0
US-10-761-006a-1_COPY_155_835 (1-681) x US-09-471-573A-2 (1-226)
QY 1 ATGGAGAACACATCAGGATTCCTAGGACCCCTCGTGTGTACAGCGGGGTTTTCTC 60
Db 1 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20
QY 61 TCGTTGACAGGATCCTCACATACCGCAGAGTCTAGACTC-TGGTGGACTTCTCTCAAT 119
```



```
Alignment Scores:
Pred. No.: 5,22e-117 Length: 226
Score: 1163.00 Matches: 208
Percent Similarity: 96.04% Conservative: 10
Best Local Similarity: 91.63% Mismatches: 8
Query Match: 91.00% Indels: 2
DB: 6 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x 5196194-21 (1-226)

QY 1 ATGGAGAACACATCAGGATTCCTAGAGACCCCTGCTGCTGTACAGCGGGGTTTTTC 60
Db 1 MetGluAsnIleThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20
QY 61 TCGTTGACAGAATCCTCACAATACCGCAGAGCTTAGACTC-TGGTGGAGCTTCTCTCAAT 119
Db 21 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 40
QY 120 TTTCTAGGGAGACACCCAGGTCTTCTGCGCAAAATTCGAGTCCCACTCCCAATCA 179
Db 41 PheLeuGlySerProValCysLeu-GlyGlnAsnSerGlnSerProThrSerAsnHi 60
QY 180 CTCACCAACCTCTGTCTCCCAATTTGCTGCTGCTATCGCTGGATGTCGCGGGTTT 239
Db 60 sSerProThrSerCysProProlleCysProGlyTyrArgTrpMetCysLeuArgArgPh 80
QY 240 TATCATATTCCTCTTCATCTGCTGCTATGCTCTCATCTTCTTGTGTTCTTCTGACTA 299
Db 80 eIleIlePheLeuPheIleLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTy 100
QY 300 CCAAGGTATGTCGCCGTTGCTCTACTTCCATCCATCATCTCGGCTTTCGCAAG 359
Db 100 rGlnGlyMetLeuProValCysProLeuIleProGlySerThrThrSerThrGlyPr 120
QY 360 ATCGAGACCTGACGACTCTCTCAAGGAACTCTACGTTTCCCTCTGTTGCTGTAC 419
Db 120 oCysLysThrCysThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysTh 140
QY 420 AAAACCTTCGAGCGGAACTGCTATGTTATTCCTCCATCCATCATCTCGGCTTTCGCAAG 479
Db 140 rLysProThrAspGlyAsnCysThrCysIleProlleProSerSerTrpAlaPheAlaTy 160
QY 480 ATTCTATGGAGTGGGCTCAGTCCGTTTCTGCTGCTCAGTCCCTTCTAGTCCATTTGT 539
Db 160 sTyrLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValProPheVa 180
QY 540 TCAGTGGTTCGTAGGGCTTCCCACTGTTGCTTTCAGTTATATGATGATGTGTA 599
Db 180 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerAlaIleTrpMetMetTrpTy 200
QY 600 TTGGGGCGGAGTCTGTACAACATCTTGAGTCCCTTTTACCTCTATTACCAATTTCTT 659
Db 200 rTrpGlyProSerLeuTySerIleValSerProPheIleProLeuLeuProIlePhePh 220
QY 660 TTGCTTTGGGTATACATT 678
Db 220 eCysLeuTrpValTyIle 226
```

```
RESULT 11
5436139-5
; Patent No. 5436139
; APPLICANT: RUTTER, WILLIAM J.;GOODMAN, HOWARD M.
; TITLE OF INVENTION: NON-PASSAGEABLE VIRUS
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/058/89,993
; FILING DATE: 08-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 571,331
; FILING DATE: 22-AUG-1990
; APPLICATION NUMBER: 513,055
; FILING DATE: 12-JUL-1983
; APPLICATION NUMBER: 107,267
```

```
; FILING DATE: 21-DEC-1979
; APPLICATION NUMBER: 41,909
; FILING DATE: 24-MAY-1979
; SEQ ID NO:5:
; LENGTH: 226
5436139-5

Alignment Scores:
Pred. No.: 5,22e-117 Length: 226
Score: 1163.00 Matches: 208
Percent Similarity: 96.04% Conservative: 10
Best Local Similarity: 91.63% Mismatches: 8
Query Match: 91.00% Indels: 2
DB: 6 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x 5436139-5 (1-226)

QY 1 ATGGAGAACACATCAGGATTCCTAGAGACCCCTGCTGCTGTACAGCGGGGTTTTTC 60
Db 1 MetGluAsnIleThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20
QY 61 TCGTTGACAGAATCCTCACAATACCGCAGAGCTTAGACTC-TGGTGGAGCTTCTCTCAAT 119
Db 21 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 40
QY 120 TTTCTAGGGAGACACCCAGGTTCCTGCGCAAAATTCGCAAGTCCCACTCCCAATCA 179
Db 41 PheLeuGlySerProValCysLeu-GlyGlnAsnSerGlnSerProThrSerAsnHi 60
QY 180 CTCACCAACCTCTGTCTCCCAATTTGCTGCTGCTATCGCTGGATGTCGCGGGTTT 239
Db 60 sSerProThrSerCysProProlleCysProGlyTyrArgTrpMetCysLeuArgArgPh 80
QY 240 TATCATATTCCTCTTCATCTGCTGCTATGCTCTCATCTTCTTGTGTTCTTCTGACTA 299
Db 80 eIleIlePheLeuPheIleLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTy 100
QY 300 CCAAGGTATGTCGCCGTTGCTCTACTTCCAGGAACTCAACACACGACGCGGGCC 359
Db 100 rGlnGlyMetLeuProValCysProLeuIleProGlySerThrThrSerThrGlyPr 120
QY 360 ATCGAGACCTGACGACTCTCTCAAGGAACTCTAGTTCCTCTGTTGCTGTAC 419
Db 120 oCysLysThrCysThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysTh 140
QY 420 AAAACCTTCGAGCGGAACTGCTATGTTATTCCTCCATCCATCATCTCGGCTTTCGCAAG 479
Db 140 rLysProThrAspGlyAsnCysThrCysIleProlleProSerSerTrpAlaPheAlaTy 160
QY 480 ATTCTATGGAGTGGGCTCAGTCCGTTTCTGCTGCTCAGTCCCTTCTAGTCCATTTGT 539
Db 160 sTyrLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValProPheVa 180
QY 540 TCAGTGGTTCGTAGGGCTTCCCACTGTTGCTTTCAGTTATATGATGATGTGTA 599
Db 180 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerAlaIleTrpMetMetTrpTy 200
QY 600 TTGGGGCGGAGTCTGTACAACATCTTGAGTCCCTTTTACCTCTATTACCAATTTCTT 659
Db 200 rTrpGlyProSerLeuTySerIleValSerProPheIleProLeuLeuProIlePhePh 220
QY 660 TTGCTTTGGGTATACATT 678
Db 220 eCysLeuTrpValTyIle 226

RESULT 12
US-09-721-480-5
; Sequence 5, Application US/09721480
; Patent No. 6740323
; GENERAL INFORMATION:
; APPLICANT: Selby, Mark
; APPLICANT: Glazer, Edward
; APPLICANT: Houghton, Michael
```

```
; TITLE OF INVENTION: HBV/HCV VIRUS-LIKE PARTICLE
; FILE REFERENCE: PP01635.002
; CURRENT APPLICATION NUMBER: US/09/721.480
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: pCMVII opti
; OTHER INFORMATION: 330 E1/SAG
US-09-721-480-5

Alignment Scores:
Pred. No.: 6.44e-117 Length: 390
Score: 1163.00 Matches: 208
Percent Similarity: 96.04% Conservative: 10
Best Local Similarity: 91.63% Mismatches: 8
Query Match: 91.00% Indels: 2
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x US-09-721-480-5 (1-390)
Qy 1 ATGGAGAACACACATCAGGATTCCTAGGACCCCTGCTCGTGTACAGGGGGGTTTTTC 60
Db 165 MetGluAsnIleThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 184
Qy 61 TCGTTGACAGAATCTCACAATACCGCAGAGTCTAGACTC-TGGTGGACTTCTCTCAAT 119
Db 185 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 204
Qy 120 TTTCTAGGGGAGACACCCAGCTGTCTCGGCCAAAATTCGCAGTCCCAACCTCCAATCA 179
Db 205 PheLeuGlyGlySerProValCysLeu-GlyGlnAsnSerGlnSerProThrSerAsnH1 224
Qy 180 CTCACCAACCTCTTGTCTCCAAATTTGTCCTGGCTATCGCTGGAGTGTCTGGCGGCTTT 239
Db 224 sSerProThrSerCysProProIleCysProGlyTyArgTrpMetCysLeuArgGPh 244
Qy 240 TATCATATCTCTTCATCCTGCTATGCTATGCTCATCTTCTGTTGTTCTTCTGGACTA 299
Db 244 eIleIlePheLeuPheIleLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTy 264
Qy 300 CCAGGTATGTGCGCGTTTGTCTCTACTTCCAGGAACATCAACCCAGCAGCGGGGCC 359
Db 264 rGlnGlyMetLeuProValCysProIleProGlySerThrThrThrSerThrGlyPr 284
Qy 360 ATGCAAGACCTGCAGACTCTCGCTCAAGGAAACTCTACGTTTCCCTCTTCTGCTGTAC 419
Db 284 oCysIleThrCysThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysTh 304
Qy 420 AAAACCTTCGGACGGAACATGCACTGTATTTCCCATCCCATCATCTCGGCTTTCGCAAG 479
Db 304 rLysProThrAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaLy 324
Qy 480 ATTCTATGGAGTGGGCTCAGTCCGTTTCTCGGCTCAGTTTACTAGTGCCTATTTGT 539
Db 324 sTyrLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa 344
Qy 540 TCAGTGGTTCGTAGGGCTTTCCCCACTGTTTGGCTTTTCAGTTATATGATGATGTGTA 599
Db 344 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerAlaIleTrpMetMetTrpTy 364
Qy 600 TTGGGGCGAAGTCTGTACAACTCTGAGTCCCTTTTACTCTATTACCAATTTCTT 659
Db 364 rTrpGlyProSerLeuTySerIleValSerProPheIleProLeuLeuProIlePhePhe 384
Qy 660 TTGCTTTGGGTATACATT 678
Db 384 eCysLeuTrpValTyIle 390
```

```
RESULT 13
US-09-721-480-7
; Sequence 7, Application US/09721480
; Patent No. 6740323
; GENERAL INFORMATION:
; APPLICANT: Selby, Mark
; APPLICANT: Glazer, Edward
; APPLICANT: Houghton, Michael
; TITLE OF INVENTION: HBV/HCV VIRUS-LIKE PARTICLE
; FILE REFERENCE: PP01635.002
; CURRENT APPLICATION NUMBER: US/09/721.480
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: plasmid
; OTHER INFORMATION: pCMV-II-E2661-sAg
US-09-721-480-7

Alignment Scores:
Pred. No.: 7.24e-117 Length: 531
Score: 1163.00 Matches: 208
Percent Similarity: 96.04% Conservative: 10
Best Local Similarity: 91.63% Mismatches: 8
Query Match: 91.00% Indels: 2
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x US-09-721-480-7 (1-531)
Qy 1 ATGGAGAACACACATCAGGATTCCTAGGACCCCTGCTCGTGTACAGGGGGGTTTTTC 60
Db 306 MetGluAsnIleThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 325
Qy 61 TCGTTGACAGAATCTCACAATACCGCAGAGTCTAGACTC-TGGTGGACTTCTCTCAAT 119
Db 326 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 345
Qy 120 TTTCTAGGGGAGACACCCAGCTGTCTCGGCCAAAATTCGCAGTCCCAACCTCCAATCA 179
Db 346 PheLeuGlyGlySerProValCysLeu-GlyGlnAsnSerGlnSerProThrSerAsnH1 365
Qy 180 CTCACCAACCTCTTGTCTCCAAATTTGTCCTGGCTATCGCTGGAGTGTCTGGCGGCTTT 239
Db 365 sSerProThrSerCysProProIleCysProGlyTyArgTrpMetCysLeuArgGPh 385
Qy 240 TATCATATCTCTTCATCCTGCTATGCTATGCTCATCTTCTGTTGTTCTTCTGGACTA 299
Db 385 eIleIlePheLeuPheIleLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTy 405
Qy 300 CCAGGTATGTGCGCGTTTGTCTCTACTTCCAGGAACATCAACCCAGCAGCGGGGCC 359
Db 405 rGlnGlyMetLeuProValCysProIleProGlySerThrThrThrSerThrGlyPr 425
Qy 360 ATGCAAGACCTGCAGACTCTCGCTCAAGGAAACTCTACGTTTCCCTCTTCTGCTGTAC 419
Db 425 oCysIleThrCysThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysTh 445
Qy 420 AAAACCTTCGGACGGAACATGCACTGTATTTCCCATCCCATCATCTCGGCTTTCGCAAG 479
Db 445 rLysProThrAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaLy 465
Qy 480 ATTCTATGGAGTGGGCTCAGTCCGTTTCTCGGCTCAGTTTACTAGTGCCTATTTGT 539
Db 465 sTyrLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa 485
Qy 540 TCAGTGGTTCGTAGGGCTTTCCCCACTGTTTGGCTTTTCAGTTATATGATGATGTGTA 599
Db 485 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerAlaIleTrpMetMetTrpTy 505
```


; TELEFAX: 610-270-5090

; TELEX:

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 424 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; US-08-760-797A-3

Alignment Scores:

Pred. No.:	1,48-116	Length:	424
Score:	1160.00	Matches:	207
Percent Similarity:	96.04%	Conservative:	11
Best Local Similarity:	91.19%	Mismatches:	8
Query Match:	90.77%	Indels:	2
DB:	1	Gaps:	0

US-10-761-006A-1_COPY_155_835 (1-681) x US-08-760-797A-3 (1-424)

Qy	1	ATGGAGAACACAAATCAGAGATTCCTAGGACCCCTGCTCGTGTTCACAGCGGGGTTTTTC	60
Db	199	MetGluAsnIleThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe	218
Qy	61	TCGTTGACAGAAATCCTCAATACCGCAGAGTCTAGACTC-TGGTGGACTTCTCTCAAT	119
Db	219	LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn	238
Qy	120	TTTCTAGGGGAGCACCACCGTGTCTCTGCGCCAAATTCGCAGTCCCAACCTCCCAATCA	179
Db	239	PheLeuGlyGlySerProValCysLeu-GlyGlnAsnSerGlnSerProThrSerAsnHi	258
Qy	180	CTACCAACCTCTTGTCTCCAAATTTGTCCTGGCTATCGCTGGAGTGTCTGGGCGCTTT	239
Db	258	sSerProThrSerCysProIleCysProGlyTyArgTrpMetCysLeuArgArgPh	278
Qy	240	TATCATATTCTCTTCATCTCTGCTGCTATGCCTCATCTTCTTGTGGTCTTCTGGACTA	299
Db	278	eIlePheLeuPheIleLeuLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTy	298
Qy	300	CCAGGTATGTTGGCGGTTTGTCTCTCTACTTCCAGGAACATCAACACACGACGCGGCC	359
Db	298	rGlnGlyMetLeuProValCysProLeuIleProGlySerThrThrAsnThrGlyPr	318
Qy	360	ATGCAAGACTGACGACTCTCTGCTCAAGAAACTCTACGTTTCCCTCTTGTGCTGTAC	419
Db	318	oCysLysThrCysThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysTh	338
Qy	420	AAACCTTCGGACGGAACCTGCACTTGTATTCCCATCCCATCATCTGGGCTTTCGCAAG	479
Db	338	rLysProThrAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaLy	358
Qy	480	ATTCTATGGAGTGGGCTCAGTCCGTTTCTCTGCTCAGTTTACTAGTGCATTTGT	539
Db	358	sTyLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa	378
Qy	540	TCAGTGTTCGTAGGGCTTCCCCACCTGTTTGGCTTTCAGTTATATGATGATGTGGA	599
Db	378	lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerAlaIleTrpMetMetTrpTy	398
Qy	600	TTGGGGGGAGTCTGTACACATCTTGAGTCCCTTTTACCTCTATATACCAATTTCTT	659
Db	398	rTrpGlyProSerLeuTySerIleValSerProPheIleProLeuLeuProIlePhePh	418
Qy	660	TTGTCTTTGGGTATACATT	678
Db	418	eCysLeuTrpValIleIle	424

Search completed: December 27, 2005, 21:03:41

Job time : 52.4 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

QM nucleic - protein search, using frame_plus_n2p model
Run on: December 27, 2005, 20:41:12 ; Search time 227.454 Seconds
(without alignments)
4224.720 Million cell updates/sec

Title: US-10-761-006A-1_COPY_155_835
Perfect score: 1278
Sequence: 1 ATGAGACACACATCAGG.....GTCTTGGGTATACATTAA 681

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2166443 seqs, 705528306 residues
Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlp
-Q/cgn2_1/USPTO.spool_p/US10761006/runat_27122005_192900_15751/app_query.fasta_1.1102
-DB=uniprot -OPMT=fastan -SUFFIX=rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORW=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USRR=US10761006 @CNG 1.1 580 @runat_27122005_192900_15751 -ICPU=3
-NO.MMAP -LARGESQUERY -NEG.SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Uniprot_05.80: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1220	95.5	226	2	Q9DUJ4_HPBVO Q9duj4 hepatitis b
2	1213	94.9	400	2	Q7TDS6_HPBVO Q7tde6 hepatitis b
3	1212	94.8	226	2	Q9DHA2_HPBVO Q9dha2 hepatitis b
4	1210	94.7	226	2	Q80GU3_HPBVO Q80gu3 hepatitis b
5	1210	94.7	400	2	Q80GU5_HPBVO Q80gu5 hepatitis b
6	1209	94.6	400	2	Q4FD52_HPBVO Q4fd52 hepatitis b
7	1208	94.5	400	2	Q5SDK8_HPBVO Q5sdk8 hepatitis b
8	1208	94.5	400	2	Q91GW8_HPBVO Q91gw8 hepatitis b
9	1207	94.4	226	2	Q9PX13_HPBVO Q9px13 hepatitis b
10	1207	94.4	392	2	Q9QC36_HPBVR Q9qc36 hepatitis b
11	1206	94.4	226	2	Q9W966_HPBVO Q9w966 hepatitis b
12	1206	94.4	226	2	Q9YK19_HPBVO Q9yk19 hepatitis b
13	1206	94.4	399	1	P03140 hepatitis b
14	1206	94.4	400	2	Q68RP7_HPBVO Q68rp7 hepatitis b
15	1206	94.4	400	2	Q7TDR4_HPBVO Q7tdr4 hepatitis b
16	1206	94.4	400	2	Q80MQ7_HPBVO Q80mq7 hepatitis b

17	1206	94.4	400	2	Q8VIM3_HPBVO Q8vim3 hepatitis b
18	1206	94.4	400	2	Q91GW7_HPBVO Q91gw7 hepatitis b
19	1206	94.4	400	2	Q9Y2T1_HPBVO Q9y2t1 hepatitis b
20	1205	94.3	226	2	Q9DUJ3_HPBVO Q9duj3 hepatitis b
21	1204	94.2	226	2	Q9DUJ6_HPBVO Q9duj6 hepatitis b
22	1204	94.2	226	2	Q9Q3D1_HPBVO Q9q3d1 hepatitis b
23	1204	94.2	226	2	Q9QT41_HPBVO Q9qt41 hepatitis b
24	1204	94.2	400	2	Q91590_HPBVO Q91590 hepatitis b
25	1204	94.2	400	2	Q7TDR2_HPBVO Q7tdr2 hepatitis b
26	1203	94.1	226	2	Q5KR14_HPBVO Q5kr14 hepatitis b
27	1203	94.1	226	2	Q9DHB0_HPBVO Q9dho0 hepatitis b
28	1203	94.1	226	2	Q9DUK0_HPBVO Q9duko hepatitis b
29	1203	94.1	226	2	Q9PX64_HPBVO Q9px64 hepatitis b
30	1203	94.1	281	2	Q80GU4_HPBVO Q80gu4 hepatitis b
31	1203	94.1	400	2	Q913B0_HPBVO Q913b0 hepatitis b
32	1203	94.1	400	2	Q9DTC8_HPBVO Q9dte8 hepatitis b
33	1202	94.1	226	2	Q80GX9_HPBVO Q80gx9 hepatitis b
34	1202	94.1	226	2	Q9DH82_HPBVO Q9dh82 hepatitis b
35	1202	94.1	226	2	Q9DUJ1_HPBVO Q9duj1 hepatitis b
36	1202	94.1	226	2	Q9QT46_HPBVO Q9qt46 hepatitis b
37	1202	94.1	226	2	Q9WLL2_HPBVO Q9wll2 hepatitis b
38	1202	94.1	400	2	Q72030_HPBVO Q72030 hepatitis b
39	1202	94.1	400	2	Q80GY1_HPBVO Q80gy1 hepatitis b
40	1202	94.1	400	2	Q80MR0_HPBVO Q80mr0 hepatitis b
41	1202	94.1	400	2	Q91GW9_HPBVO Q91gw9 hepatitis b
42	1202	94.1	400	2	Q91GX1_HPBVO Q91gx1 hepatitis b
43	1202	94.1	400	2	Q4FDF8_HPBVO Q4fd8 hepatitis b
44	1202	94.1	400	2	Q4FDF8_HPBVO Q4fd8 hepatitis b
45	1201	94.0	226	2	Q5G8L9_HPBVO Q5g8l9 hepatitis b

ALIGNMENTS

RESULT 1

Q9DUJ4_HPBVO
ID Q9DUJ4_HPBVO PRELIMINARY; PRT; 226 AA.
AC Q9DUJ4; 16, Created
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE Small surface polypeptide.
GN Names=OG14;
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Ogura Y., Kurosaki M., Asahina Y., Enomoto N., Marumo F., Sato C.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB029963; BAB20346.1; -; Genomic_DNA.
DR PIR; JQ2094; JQ2094.
DR PIR; JQ2095; JQ2095.
DR PIR; JQ2096; JQ2096.
DR PIR; JQ2097; JQ2097.
DR PIR; JQ2098; JQ2098.
DR PIR; JQ2099; JQ2099.
DR PIR; JQ2100; JQ2100.
DR PIR; JQ2101; JQ2101.
DR PIR; JQ2102; JQ2102.
DR PIR; JQ2106; JQ2106.
DR PIR; JQ2108; JQ2108.
DR PIR; JQ2109; JQ2109.
DR PIR; JQ2111; JQ2111.
DR PIR; JQ2112; JQ2112.
DR PIR; JQ2116; JQ2116.
DR GO; GO:0016032; P:Viral life cycle; IEA.
DR InterPro; IPR000349; Hepvir_surfaG.
DR PANTHER; PTHR10832; Hepvir_surfaG; 1.
DR Pfam; PF00695; VMSA; 1.
KW Antigen; Polypeptide.
SQ SEQUENCE 226 AA; 25375 MW; 7704565C4092E3BC CRC64;

Alignment Scores:
 Pred. No.: 9,27e-100 Length: 226
 Score: 1220.00 Matches: 222
 Percent Similarity: 98.24% Conservative: 1
 Best Local Similarity: 97.80% Mismatches: 3
 Query Match: 95.46% Indels: 0
 DB: 2 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x Q9DUJ4_HPBVO (1-226)

QY	1	ATGGAGACACACATCAGGATTCCTAGGACCCCTGCTGCTTTACAGCGGGGTTTTTC	60
DB	1	MetGluAsnThrThrSerGlyPheLeuGlyProLeuValLeuGlnAlaGlyPhePhe	20
QY	61	TGCTTTGACAAAGATCCTCAATACCGAGAGCTAGACTC-TGGTGGACTTCTCTCAAT	119
DB	21	LeuLeuThrArgGlyLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn	40
QY	120	TTTCTAGGGGAGACCCACGTTGCTGGCCAAATTCGCAGTCCCAACCTCCAATCA	179
DB	41	PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi	60
QY	180	CTCACCACTCTGCTCCCAATTTGCTGGCTATCGCTGGATGCTGCTGGGGTTTT	239
DB	60	sSerProThrSerCysProIleCysProGlyTyArgTrpMetCysLeuArgArgPh	80
QY	240	TATCATATTCCTCTTCATCTGCTGCTATGCTCTCATCTTCTTGTGGTCTTCTGACTA	299
DB	80	eillelePheLeuPheIleLeuLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTy	100
QY	300	CAAAGTATGTTGCCCGTTGCTCTACTTCCAGAAACATCAACACACGACGCGGCC	359
DB	100	rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr	120
QY	360	ATGCAAGACTCGACACTCTGCTCAAGAACTCAGTTTCCCTCTTGTGCTGTAC	419
DB	120	oCysLyThrCysThrSerProAlaGlnGlyAsnSerThrPheProSerCysCysPh	140
QY	420	AAAACTTCGGACGGAACTCGACTTGATTCCTCCATCCCATCATCTCGGGCTTCGCAAG	479
DB	140	rllysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaAr	160
QY	480	ATTCTATCGGAGTGGGCTCAGTCCGTTTCTCTGGCTCAGTTACTAGTCCCAATTGT	539
DB	160	gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValProPheVa	180
QY	540	TCAGTGTTCGTAGGGCTTTCCCACTGTTGGCTTTTCAGTTATATGATGATGCTA	599
DB	180	lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy	200
QY	600	TTGGGGGCGAAGTCTGTCAACATCTTGAGTCCCTTTTACCTCTATTACCAATTTCCT	659
DB	200	rTrpGlyProSerLeuTyAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh	220
QY	660	TTGCTCTTCGGTATACATT 678	
DB	220	eCysLeuTrpValTyIle 226	

RESULT 2
 Q7TDS6_HPBVO
 ID Q7TDS6_HPBVO PRELIMINARY; PRT; 400 AA.
 AC Q7TDS6;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Presi/preS2 surface protein.
 OS Hepatitis B virus.
 OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
 OC Orthohepadnavirus.
 OX NCBI_TaxID=10407;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Liu C.-J., Chen P.-J., Lai M.-Y., Kao J.-H., Chen D.-S.;

RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY167091; AAO41300.1; -; Genomic DNA.
 GO; GO:0016032; P:Viral life cycle; IEA.
 DR InterPro; IPR000349; Hepvir_surflag.
 DR PANTHER; PTHR10832; Hepvir_surflag; 1.
 DR Pfam; PF00695; VNSA; 1.
 KW Antigen.
 SQ SEQUENCE 400 AA; 43790 MW; 5D2PFC9D91CFE4D7 CRC64;

Alignment Scores:
 Pred. No.: 4,23e-99 Length: 400
 Score: 1213.00 Matches: 221
 Percent Similarity: 97.36% Conservative: 0
 Best Local Similarity: 97.36% Mismatches: 5
 Query Match: 94.91% Indels: 2
 DB: 2 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x Q7TDS6_HPBVO (1-400)

QY	1	ATGGAGACACACATCAGGATTCCTAGGACCCCTGCTGCTTTACAGCGGGGTTTTTC	60
DB	175	MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe	194
QY	61	TGCTTTGACAAAGATCCTCAATACCGAGAGCTAGACTC-TGGTGGACTTCTCTCAAT	119
DB	195	LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn	214
QY	120	TTTCTAGGGGAGACCCACGTTGCTGGCCAAATTCGCAGTCCCAACCTCCAATCA	179
DB	215	PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi	234
QY	180	CTCACCACTCTTCTCTCTCCAAATTTGCTGGCTATCGCTGGATGCTGCTGGCGGTTTT	239
DB	234	sSerProThrSerCysProIleCysProGlyTyArgTrpMetCysLeuArgArgPh	254
QY	240	TATCATATTCCTCTTCATCTGCTGCTATGCTCTCATCTTCTTGTGGTCTTCTGACTA	299
DB	254	eillelePheLeuPheIleLeuLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTy	274
QY	300	CAAAGTATGTTGCCCGTTGCTCTACTTCCAGAAACATCAACACACGACGCGGCC	359
DB	274	rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr	294
QY	360	ATGCAAGACTCGACACTCTGCTCAAGAACTCAGTTTCCCTCTTGTGCTGTAC	419
DB	294	oCysLyThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysPh	314
QY	420	AAAACTTCGGACGGAACTCGACTTGATTCCTCCATCCCATCATCTCGGGCTTCGCAAG	479
DB	314	rllysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaAr	334
QY	480	ATTCTATCGGAGTGGGCTCAGTCCGTTTCTCTGGCTCAGTTACTAGTCCCAATTGT	539
DB	334	gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValProPheVa	354
QY	540	TCAGTGTTCGTAGGGCTTTCCCACTGTTGGCTTTTCAGTTATATGATGATGCTA	599
DB	354	lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy	374
QY	600	TTGGGGGCGAAGTCTGTCAACATCTTGAGTCCCTTTTACCTCTATTACCAATTTCCT	659
DB	374	rTrpGlyArgSerLeuTyAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh	394
QY	660	TTGCTCTTCGGTATACATT 678	
DB	394	eCysLeuTrpValTyIle 400	

RESULT 3
 Q9DHA2_HPBVO
 ID Q9DHA2_HPBVO PRELIMINARY; PRT; 226 AA.
 AC Q9DHA2;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)

```

DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Small surface glycoprotein.
GN Name=OG37; Synonym=OG28;
OS Hepatitis B virus.
OC Viruses; Retroviridae; Orthohepadnaviruses.
OC NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Ogura Y., Kurosaki M., Ashina Y., Enomoto N., Marumo F., Sato C.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB029986; BAB20369.1; -; Genomic DNA.
DR EMBL; AB029977; BAB20360.1; -; Genomic DNA.
DR PIR; JQ2094; JQ2094.
DR PIR; JQ2095; JQ2095.
DR PIR; JQ2096; JQ2096.
DR PIR; JQ2097; JQ2097.
DR PIR; JQ2098; JQ2098.
DR PIR; JQ2099; JQ2099.
DR PIR; JQ2100; JQ2100.
DR PIR; JQ2101; JQ2101.
DR PIR; JQ2102; JQ2102.
DR PIR; JQ2106; JQ2106.
DR PIR; JQ2108; JQ2108.
DR PIR; JQ2109; JQ2109.
DR PIR; JQ2111; JQ2111.
DR PIR; JQ2112; JQ2112.
DR PIR; JQ2116; JQ2116.
DR GO; GO:0016032; P:Viral life cycle; IEA.
DR InterPro; IPR000349; Hepvir_surfaG.
DR PANTHER; PTHR10832; Hepvir_surfaG; 1.
DR Pfam; PF00695; VMSA; 1.
DR Antigen; Polyprotein.
SQ SEQUENCE 226 AA; 25406 MW; 25F97F2880749EEE CRC64;

Alignment Scores:
Pred. No.: 4,78e-99 Length: 226
Score: 1212.00 Matches: 221
Percent Similarity: 97.36% Conservative: 0
Best Local Similarity: 97.36% Mismatches: 5
Query Match: 94.84% Indels: 2
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x Q9DHA2_HPBVO (1-226)
Qy 1 ATGAGAACACAAATCAGGATCTCTAGGACCTCTCGTGTGTACAGGGGGGTTTTC 60
Db 1 MetGluasnThrThrSerGlyPheLeuGlyProLeuValLeuGlnAlaGlyPhePhe 20
Qy 61 TCGTTGACAGAACTCTCACAAATACCGCAGAGTCTAGACTC-TGGTGGACTTCTCAAT 119
Db 21 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 40
Qy 120 TTTCTAGGGGAGCACCACGTTCTCTGGCAAAATTCGAGTCCCAACTCCAATCA 179
Db 41 PheLeuGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnH1 60
Qy 180 CTCACCAACTCTGTCTCTCAATTTCTCGGTATCGGTGGATGTCTCGCGGCTTT 239
Db 60 sSerProThrSerCysProIleCysProGlyTyArgTrpMetCysLeuArgA-gPh 80
Qy 240 TATCATATTCTCTTCATCTCTGCTATGCTCATCTCTTGTGGTCTTCTGGACTA 299
Db 80 eilePheLeuPheIleLeuLeuCysLeuIlePheLeuValLeuLeuAspTy 100
Qy 300 CCAAGGTATGTTGCGCGTTTCTCTACTTTCAGGAACATCAACACACGACGGGCC 359
Db 100 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 120
Qy 360 ATCAAGACCTGACGACTCTCTGCTCAAGAACTCTACGTTTCCCTCTTGTCTGTAC 419
Db 120 oCysIsthrCysThrThrProAlaGlnGlyThrSerMetPheProSerCysCysA 140
Qy 420 AAAACCTTCGACGGAACACTGACTTGTATTCCCATCCCATCTCTGGGGCTTCGCAAG 479

RESULT 4
Q80GU3_HPBVO
ID Q80GU3_HPBVO PRELIMINARY; PRT; 226 AA.
AC Q80GU3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE S protein.
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OC NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FMU021;
RA Lin X., Zheng D.L., Xu X.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY206392; AAP06646.1; -; Genomic DNA.
DR GO; GO:0016032; P:Viral life cycle; IEA.
DR InterPro; IPR000349; Hepvir_surfaG.
DR PANTHER; PTHR10832; Hepvir_surfaG; 1.
DR Pfam; PF00695; VMSA; 1.
DR Antigen.
SQ SEQUENCE 226 AA; 25449 MW; DD79F9B6A09D5933 CRC64;

Alignment Scores:
Pred. No.: 7.2e-99 Length: 226
Score: 1210.00 Matches: 220
Percent Similarity: 97.36% Conservative: 1
Best Local Similarity: 96.92% Mismatches: 5
Query Match: 94.68% Indels: 2
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x Q80GU3_HPBVO (1-226)
Qy 1 ATGAGAACACAAATCAGGATCTCTAGGACCTCTCGTGTGTACAGGGGGGTTTTC 60
Db 1 MetGluasnThrThrSerGlyPheLeuGlyProLeuValLeuGlnAlaGlyPhePhe 20
Qy 61 TCGTTGACAGAACTCTCACAAATACCGCAGAGTCTAGACTC-TGGTGGACTTCTCAAT 119
Db 21 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 40
Qy 120 TTTCTAGGGGAGCACCACGTTCTCTGGCAAAATTCGAGTCCCAACTCCAATCA 179
Db 41 PheLeuGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnH1 60
Qy 180 CTCACCAACTCTGTCTCTCAATTTCTCGGTATCGGTGGATGTCTCGCGGCTTT 239
Db 60 sSerProThrSerCysProIleCysProGlyTyArgTrpMetCysLeuArgA-gPh 80
Qy 240 TATCATATTCTCTTCATCTCTGCTATGCTCATCTCTTGTGGTCTTCTGGACTA 299
Db 80 eilePheLeuPheIleLeuLeuCysLeuIlePheLeuValLeuLeuAspTy 100
```

QY 300 CCAAGGTATGTTGCCGCTTTGCTCTACTTCCAGGAACATCAACACACGACGCGGGCC 359
 Db 100 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrGlyPr 120
 QY 360 ATGCAAGACCTCGACACTCTGCTCAGGAAACTCTACGTTTCCCTCTTGTGTGTGTAC 419
 Db 120 oCysGlyThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysCysTh 140
 QY 420 AAAACCTTCGGACGGAACTGACACTGTTATCCCATCCCATCATCTCTGGGCTTTCCGAAG 479
 Db 140 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaLys 160
 QY 480 ATTCTATGCGAGTGGGCTCAGTCCGTTCTCTCTGGCTCAGTTACTAGTCCCAATTTGT 539
 Db 160 sPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValProPheVa 180
 QY 540 TCAGTGGTTCGAGGCTTTCCCACTGTTTGGCTTTTCCAGTTATATGATGATGTGTA 599
 Db 180 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTripy 200
 QY 600 TTGGGGGCAAGCTCTGACAACTTGTAGTCCCTTTTACCTCTATTACCAATTTCTT 659
 Db 200 rTrpGlyArgSerLeuTyAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh 220
 QY 660 TTGTCTTTGGGTATACATT 678
 Db 220 eCysLeuTrpValTyrlle 226

RESULT 5

Q80GUS_HPBVO
 ID Q80GUS_HPBVO PRELIMINARY; PRT; 400 AA.
 AC Q80GUS;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Large S protein.
 OS Hepatitis B virus.
 OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
 OC Orthohepadnavirus.
 OX NCBI_TaxID=10407;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=FMU021;
 RA Lin X., Zheng D.L., Xu X.;
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; A1206392; AAP0644.1; -; Genomic DNA.
 DR GO; GO:0016032; P:Viral life cycle; IEA.
 DR InterPro; IPR000349; Hepvir_surfaG.
 DR PANTHER; PTHR10832; Hepvir_surfaG; 1.
 DR Pfam; PF00695; VMSA; 1.
 KW Antigen.
 SQ SEQUENCE 400 AA; 43753 MW; 19ED0C82CC80792 CRC64;

Alignment Scores:
 Pred. No.: 7,82e-99 Length: 400
 Score: 1210.00 Matches: 220
 Percent Similarity: 97.36% Conservative: 1
 Best Local Similarity: 96.92% Mismatches: 5
 Query Match: 94.68% Indels: 2
 DB: 2 Gaps: 0
 US-10-761-006a-1_COPY_155_835 (1-681) x Q80GUS_HPBVO (1-400)

QY 1 ATGGAACAACAATCATCAGTTCCTAGGACCCCTGCTGTATACAGGCGGGTTTTC 60
 Db 175 MetGluAsnThrThrSerGlyPheLeuLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 194
 QY 61 TCCTTACCAAGAATCTCAATACCGAGCTAGACTC-TGGTGGAGCTTCTCTCAAT 119
 Db 195 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 214
 QY 120 TTTCTAGGGGGAGCACCCACAGTGTCTCTGCGCCAAAATTCGACGTCCCACTCCCAATCA 179

Db 215 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 234
 QY 180 CTCACCAACCTTGTCTCTCAATTTGTCCTGGCTATCGCTGGAGTGTCTCTGCGCGCTTT 239
 Db 234 sSerProThrSerCysProIleCysProGlyTyArgTrpMetCysLeuArgArgPh 254
 QY 240 TATCATATTCCTCTCTCATCTCTGCTATGCTCATCTTCTTGTGGTCTTCTTGACTA 299
 Db 254 eIleIlePheLeuPheIleLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTy 274
 QY 300 CCAGGTATGTTGCCGTTTCTCTACTTCCAGGAACATCAACACACGACGCGGGCC 359
 Db 274 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrSerThrGlyPr 294
 QY 360 ATGCAAGACCTCGACACTCTCTCTCAAGGAACTCTACGTTTCCCTCTTGTGTGTGTAC 419
 Db 294 oCysLysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysTh 314
 QY 420 AAAACCTTCGGACGGAACTGACACTTGTATTCCTCATCCCATCATCTCTGGGCTTTCCGAAG 479
 Db 314 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaLys 334
 QY 480 ATTCTATGCGAGTGGGCTCAGTCCGTTTCTCTGGCTCAGTTTACTAGTCCCAATTTGT 539
 Db 334 sPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValProPheVa 354
 QY 540 TCAGTGGTTCGAGGCTTTCCCACTGTTTGGCTTTTCCAGTTATATGATGATGTGTA 599
 Db 354 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTripy 374
 QY 600 TTGGGGGCGAAGTCTGTCAACATCTTGTAGTCCCTTTTACCTCTATTACCAATTTCTT 659
 Db 374 rTrpGlyArgSerLeuTyAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh 394
 QY 660 TTGTCTTTGGGTATACATT 678
 Db 394 eCysLeuTrpValTyrlle 400

RESULT 6

Q4FD52_HPBVO
 ID Q4FD52_HPBVO PRELIMINARY; PRT; 400 AA.
 AC Q4FD52;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Pres1/S2/S.
 GN Name=Pres1/S2/S;
 OS Hepatitis B virus.
 OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
 OC Orthohepadnavirus.
 OX NCBI_TaxID=10407;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=HK2100;
 RX Chan H.L., Tsui S.K., Tse C.H., Ng E.Y., Au T.C., Yuen L.,
 RA Bartholomeusz A., Leung K.S., Lee K.H., Locarnini S., Sung J.J.;
 RT "Epidemiological and virological characteristics of 2 subgroups of
 hepatitis B virus genotype C."
 RL J. Infect. Dis. 191:2022-2032(2005).
 DR EMBL; DQ089801; AA05311.1; -; Genomic DNA.
 SQ SEQUENCE 400 AA; 43707 MW; 2CDF00B45325BFD8 CRC64;

Alignment Scores:
 Pred. No.: 9.6e-99 Length: 400
 Score: 1209.00 Matches: 220
 Percent Similarity: 97.36% Conservative: 1
 Best Local Similarity: 96.92% Mismatches: 5
 Query Match: 94.60% Indels: 2
 DB: 2 Gaps: 0
 US-10-761-006a-1_COPY_155_835 (1-681) x Q4FD52_HPBVO (1-400)

QY 1 ATGAGAACACATCAGATTCCTAGGACCCCTGCTGTTTACAGGGGGGTTTTC 60
DB 175 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 194
QY 61 TCGTTGACAGAAATCCTCACATACCGCAGAGTCTAGACTC-TGGTGGACTTCTCTCAAT 119
DB 195 LeuLeuThrArgLeuLeuThrLeuProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 214
QY 120 TTTCTAGGGGAGACCCAGTGTCTGCGCCAAAATTCGACGTCCCAACCTCCCAATCA 179
DB 215 PheLeuGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 234
QY 180 CTCACCACTCTGCTCCCTGCTCAATATTCGCTGCTATCGCTGATGCTGCGGGTTC 239
DB 234 sSerProThrSerCysProProIleCysProGlyTyArgTrpMetCysLeuArgPh 254
QY 240 TATCATATTCCTCTTCACTGCTGCTATGCTCATCTCTGTTGGTCTCTTGGACTA 299
DB 254 eIleIlePheLeuPheLeuLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTy 274
QY 300 CCAAGGTATGTCCTGCTGCTCAATATTCGCTGCTATCGCTGATGCTGCGGGTTC 359
DB 274 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 294
QY 360 ATGACAGACCTGCTGCTGCTCAAGAACTCTAGTTCCTGCTGCTGCTGCTGCTGCT 419
DB 294 oCysLeuThrCysThrProAlaGlnGlyThrSerMetPheProSerCysCysTh 314
QY 420 AAAACCTTCGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 479
DB 314 rLysProSerAspGlyAsnCysThrCysIleProIleProSerProPheLeuProLeuPhePh 394
QY 660 TTGCTCTTGGGTATACATT 678
DB 394 eCysLeuTrpValTyrIle 400

RESULT 7

Q5SDK8 HPBV0
ID Q5SDK8 HPBV0 PRELIMINARY; PRT; 400 AA.
AC Q5SDK8;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Envelope protein.
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OK NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=He100;
RA Song B.-C., Kim H., Kim S.-H., Cha C.-Y., Kook Y.-H., Kim B.-J.;
RT "Comparison of full length sequences of hepatitis B virus isolates in
RT hepatocellular carcinoma patients and asymptomatic carriers of
RT Korea."
RL J. Med. Virol. 75:13-19(2005).
DR EMBL; AY641560; AAV52022.1; -; Genomic DNA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0016032; P:viral life cycle; IEA.
DR InterPro; IPR000349; Hepvir_surfaG.

DR PANTHER; PTHR10832; Hepvir_surfaG; 1.
DR Pfam; PF00695; yMSA; 1
KW Antigen; Envelope protein.
SQ SEQUENCE 400 AA; 43712 MW; 5C02A7936484BAC1 CRC64;

Alignment Scores:

Pred. No.: 1,18e-98 Length: 400
Score: 1208.00 Matches: 220
Percent Similarity: 96.92% Conservative: 0
Best Local Similarity: 96.92% Mismatches: 6
Query Match: 94.52% Indels: 2
DB: 2 Gaps: 0

US-10-761-006a-1_COPY_155_835 (1-681) x Q5SDK8_HPVB0 (1-400)

QY 1 ATGAGAACACATCAGATTCCTAGGACCCCTGCTGTTTACAGGGGGGTTTTC 60
DB 175 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 194
QY 61 TCGTTGACAGAAATCCTCACATACCGCAGAGTCTAGACTC-TGGTGGACTTCTCTCAAT 119
DB 195 LeuLeuThrArgLeuLeuThrLeuProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 214
QY 120 TTTCTAGGGGAGACCCAGTGTCTGCGCCAAAATTCGACGTCCCAACCTCCCAATCA 179
DB 215 PheLeuGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 234
QY 180 CTCACCACTCTGCTCCCTGCTCAATATTCGCTGCTATCGCTGATGCTGCGGGTTC 239
DB 234 sSerProThrSerCysProProIleCysProGlyTyArgTrpMetCysLeuArgPh 254
QY 240 TATCATATTCCTCTTCACTGCTGCTATGCTCATCTCTGTTGGTCTCTTGGACTA 299
DB 254 eIleIlePheLeuPheLeuLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTy 274
QY 300 CCAAGGTATGTCCTGCTGCTCAATATTCGCTGCTATCGCTGATGCTGCGGGTTC 359
DB 274 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 294
QY 360 ATGACAGACCTGCTGCTGCTCAAGAACTCTAGTTCCTGCTGCTGCTGCTGCTGCT 419
DB 294 oCysLeuThrCysThrProAlaGlnGlyThrSerMetPheProSerCysCysTh 314
QY 420 AAAACCTTCGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 479
DB 314 rLysProSerAspGlyAsnCysThrCysIleProIleProSerProPheLeuProLeuPhePh 334
QY 480 ATTCTCTATGCGAGTGGGCTCAGTCCGTTTCTGCTGCTGCTGCTGCTGCTGCTGCT 539
DB 334 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValProPheVa 354
QY 540 TCAGTGTTCGTAGGGGCTTCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 599
DB 354 lGlnTrpPheAlaGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTTrp 374
QY 600 TTGGGGGCGAAGTCTGTACAACTCTTGTAGTCCCTTTTACCTCTATTACCAATTTCTT 659
DB 374 rTrpGlyProSerLeuTyAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh 394
QY 660 TTGCTCTTGGGTATACATT 678
DB 394 eCysLeuTrpValTyrIle 400

RESULT 8

Q91GW8 HPBV0
ID Q91GW8 HPBV0 PRELIMINARY; PRT; 400 AA.
AC Q91GW8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE HBaG.
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;

```
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1] NUCLEOTIDE SEQUENCE.
RP Shi S., Dong J., Huangfu J.;
RA Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AY040799; AAK94658.1; -; Genomic_DNA.
DR PIR; JQ2094; JQ2094.
DR PIR; JQ2095; JQ2095.
DR PIR; JQ2096; JQ2096.
DR PIR; JQ2097; JQ2097.
DR PIR; JQ2098; JQ2098.
DR PIR; JQ2099; JQ2099.
DR PIR; JQ2100; JQ2100.
DR PIR; JQ2101; JQ2101.
DR PIR; JQ2102; JQ2102.
DR PIR; JQ2106; JQ2106.
DR PIR; JQ2108; JQ2108.
DR PIR; JQ2109; JQ2109.
DR PIR; JQ2111; JQ2111.
DR PIR; JQ2112; JQ2112.
DR PIR; JQ2116; JQ2116.
DR GO; GO:0016032; P:Viral life cycle; IEA.
DR InterPro; IPR000349; Hepvir_surfaG.
DR PANTHER; PTHR10832; Hepvir_surfaG; 1.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.
SQ SEQUENCE 400 AA; 43682 MW; 283B506EE6157CAF CRC64;

Alignment Scores:
Pred. No.: 1,18e-98 Length: 400
Score: 1208.00 Matches: 220
Percent Similarity: 97.36% Conservative: 1
Best Local Similarity: 96.92% Mismatches: 5
Query Match: 94.52% Indels: 2
DB: 2 Gaps: 0

US-10-761-006a-1_copy_155_835 (1-681) x Q91GW8_HPBVO (1-400)

QY 1 ATGGAGAACACATCATCGATTCTAGGACCCCTGCTGTTTACAGCGGGGTTTTC 60
DB 175 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 194

QY 61 TCGTTCACAGAAATCTCTCAATACCGCAGAGCTAGACTC-TGGTGGAGCTTCTCTCAAT 119
DB 195 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpThrSerLeuAsn 214

QY 120 TTTCTAGGGGAGACCCACGTTTCTCGGCCAAAATTCGAGTCCCAACCTCCCAATCA 179
DB 215 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 234

QY 180 CTCACCAACTCTTGTCCTCAATTTGTCCTGCTATCGCTGATGTGTCGGGGGTTT 239
DB 234 sSerProThrSerCysProProIleCysProGlyTyArgTrpMetCysLeuArgArgPh 254

QY 240 TATCATATTCTTCTATCTGCTGCTATGCTGCTATCTTCTTCTGTTCTTCTGACTA 299
DB 254 elleIlePheLeuPheIleLeuLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTy 274

QY 300 CCAAGTATGTTGTCGGCTTGTCTCTACTTCCAGGAACATCAACACGACGACGGGCC 359
DB 274 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 294

QY 360 ATGCAGACCTGACGACTCTGCTCTCAGAAACTCTAGTTCCCTCTTGTCTGTGTAC 419
DB 294 oCysIleThrCysThrSerProAlaGlnGlyThrSerMetPheProSerCysCysTh 314

QY 420 AAAACCTTCGACGGAAACTGCACTTGATTCCATCCCATCTCTGGGCTTTCGCAAG 479
DB 314 rIlySProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaAr 334

QY 480 ATTCTATGGAGTGGGCCCTCAGTCGCTTCTCTCGCTCAGTTTACTAGTCCATTGTT 539
DB 480 ATTCGAGAACACATCATCAGGATTCCTAGGACCCCTCTCTGTTTACAGCGGGGTTTTC 60

334 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa 354
540 TCAGTGGTTCGTAGGGCTTTCCCCACACTGTTGGCTTTTTCAGTTATATGATGATCGGTA 599
354 iGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 374
600 TTGGGGCGCAAGTCTGTCAACATCTTCAGTCCCTTTTACCTCTATTACCAATTTCTT 659
374 rTrpGlyProSerLeuTyAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh 394
660 TTGCTTTGGGTATACATT 678
394 eCysLeuTrpValTyIle 400

RESULT 9
Q9PX13_HPBVO
ID Q9PX13_HPBVO PRELIMINARY; PRT; 226 AA.
AC Q9PX13;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-FEB-2005 (TremBLrel. 29, Last annotation update)
DE Surface antigen.
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OC NCBI_TaxID=10407;
RN [1] NUCLEOTIDE SEQUENCE.
RP STRAIN=genotype C;
RX MEDLINE=20528425; PubMed=11074476;
RX DOI=10.1002/1096-9071(200012)62:4<471::AID-JMV12>3.3.CO;2-O;
RA Rokuhara A., Tanaka E., Yagi S., Mizokami M., Hashikura Y.,
RA Kawasaki S., Kiyosawa K.;
RT "De novo infection of hepatitis B virus in patients with orthotopic
RT liver transplantation: analysis by determining complete sequence of
RT the genome.";
RL J. Med. Virol. 62:471-478(2000).
DR EMBL; AB030517; BAA82654.1; -; Genomic DNA.
DR EMBL; AB030516; BAA82653.1; -; Genomic_DNA.
DR PIR; JQ2094; JQ2094.
DR PIR; JQ2095; JQ2095.
DR PIR; JQ2096; JQ2096.
DR PIR; JQ2097; JQ2097.
DR PIR; JQ2098; JQ2098.
DR PIR; JQ2099; JQ2099.
DR PIR; JQ2100; JQ2100.
DR PIR; JQ2101; JQ2101.
DR PIR; JQ2102; JQ2102.
DR PIR; JQ2106; JQ2106.
DR PIR; JQ2108; JQ2108.
DR PIR; JQ2109; JQ2109.
DR PIR; JQ2111; JQ2111.
DR PIR; JQ2112; JQ2112.
DR PIR; JQ2116; JQ2116.
DR GO; GO:0016032; P:Viral life cycle; IEA.
DR InterPro; IPR000349; Hepvir_surfaG.
DR PANTHER; PTHR10832; Hepvir_surfaG; 1.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.
SQ SEQUENCE 226 AA; 25394 MW; 835E71D8909A9C21 CRC64;

Alignment Scores:
Pred. No.: 1,33e-98 Length: 226
Score: 1207.00 Matches: 220
Percent Similarity: 96.92% Conservative: 0
Best Local Similarity: 96.92% Mismatches: 6
Query Match: 94.44% Indels: 2
DB: 2 Gaps: 0

US-10-761-006a-1_copy_155_835 (1-681) x Q9PX13_HPBVO (1-226)

QY 1 ATGGAGAACACATCATCAGGATTCCTAGGACCCCTCTCTGTTTACAGCGGGGTTTTC 60
DB 1 ATTCGAGAACACATCATCAGGATTCCTAGGACCCCTCTCTGTTTACAGCGGGGTTTTC 60
```


Db 1 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20
 QY 61 TCCTTGACAGAAATCTCAATACCGAGAGCTAGACTC-TGGTGGAGCTTCTCTCAAT 119
 Db 21 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAen 40
 QY 120 TTTCTAGGGGAGCACCACGCTGTTCTCGCCAAATTCGCGAGTCCCAACCTCCCAATCA 179
 Db 41 PheLeuGlyAlaProThrCys-ProGlyGlnAenSerGlnSerProThrSerAenH 60
 QY 180 CTCACCAACCTCTTGCTCCCAATTTGTCTCGCTATCGTGTGATGTGTGGGGGTTT 239
 Db 60 sSerProThrSerCysProProThrCysProGlyTyArgTrpMetCysLeuArgArgPh 80
 QY 240 TATCATATCTCTTTCATCTGCTGCTATGCTATGCTATGCTATGCTATGCTATGCTATG 299
 Db 80 eIleIlePheLeuPheIleLeuLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTy 100
 QY 300 CCAAGGTATGTTGCCGTTTGTCTCTACTTCCAGGAACATCAACACACGACGCGGGCC 359
 Db 100 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrGlyPr 120
 QY 360 ATGCAAGACCTGCACGACTCTCTCTCAAGGAAACTCTAGCTTTCCCTCTTGTGTGTATC 419
 Db 120 oCysIysThrCysThrThrProAlaGlnGlyThrSerMetPheProSerCysCysTh 140
 QY 420 AAAACCTTCGACGGAACCTGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATG 479
 Db 140 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAla 160
 QY 480 ATTCTATGGAGTGGAGTGGCTCAGTGGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539
 Db 160 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValProPheVa 180
 QY 540 TCAGTGTGCTGAGGGCTTTCCCTCCACTGTTGGCTTTTCAGTTATGATGATGATGATGATG 599
 Db 180 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 200
 QY 600 TTGGGGGGGAGTGTGTACAACTGTGAGTCCCTTTTACCTCTATTACCAATTTCTT 659
 Db 200 rTrpGlyProSerLeuTyAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh 220
 QY 660 TTGCTCTTTGGGTATACATT 678
 Db 220 eCysLeuTrpValTyIle 226

RESULT 10

Q9QC26_HPBVR
 ID Q9QC26_HPBVR PRELIMINARY; PRT; 392 AA.
 AC Q9QC26;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Large envelope protein.
 OS Hepatitis B virus (subtype adr).
 OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
 OC Orthohepadnavirus.
 OX NCBI_TaxID=106820;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Dong J., Cheng J., Wang Q.H., Liu Y.Z., Xia X.B., Song H.B., Wang G.,
 RA Zhong Y.W., Shi S.S., Yang J.Z., Liu Y.;
 RA Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AF191303; AAF05722.2; -; Genomic_DNA.
 DR PIR; JQ2094; JQ2094.
 DR PIR; JQ2095; JQ2095.
 DR PIR; JQ2096; JQ2096.
 DR PIR; JQ2097; JQ2097.
 DR PIR; JQ2098; JQ2098.
 DR PIR; JQ2099; JQ2099.
 DR PIR; JQ2100; JQ2100.
 DR PIR; JQ2101; JQ2101.
 DR PIR; JQ2102; JQ2102.

DR PIR; JQ2106; JQ2106.
 DR PIR; JQ2108; JQ2108.
 DR PIR; JQ2109; JQ2109.
 DR PIR; JQ2111; JQ2111.
 DR PIR; JQ2112; JQ2112.
 DR PIR; JQ2116; JQ2116.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0016032; P:viral life cycle; IEA.
 DR InterPro; IPR000349; Hepvir_surfac.
 DR PANTHER; PTHR10832; Hepvir_surfac; 1.
 DR Pfam; PF00695; vmsa; 1.
 KW Antigen; Envelope protein.
 SQ SEQUENCE 392 AA; 42699 MW; 0394CD9EFCAL1757 CRC64;

Alignment Scores:

Pred. No.: 1,44e-98 Length: 392
 Score: 1207.00 Matches: 220
 Percent Similarity: 96.92% Conservative: 0
 Best Local Similarity: 96.92% Mismatches: 6
 Query Match: 94.44% Indels: 2
 DB: 2 Gaps: 0

US-10-761-006a-1_COPY_155_835 (1-681) x Q9QC26_HPBVR (1-392)

QY 1 ATGGAGAACACAACTCAGGATTCCTAGGACCCCTGCTGCTGTACAGCGGGGTTTTC 60
 Db 167 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 186
 QY 61 TCGTTGACAGAAATCTCAATACCGAGAGCTAGACTC-TGGTGGAGCTTCTCTCAAT 119
 Db 187 SerLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAen 206
 QY 120 TTTCTAGGGGAGCACCACGCTGTTCTCGCCAAATTCGCGAGTCCCAACCTCCCAATCA 179
 Db 207 PheLeuGlyAlaProThrCys-ProGlyGlnAenSerGlnSerProThrSerAenH 226
 QY 180 CTCACCAACCTCTTGCTCCCAATTTGTCTCGCTATGCTGCTGCTGCTGCTGCTGCTGCT 239
 Db 226 sSerProThrSerCysProProThrCysProGlyTyArgTrpMetCysLeuArgArgPh 246
 QY 240 TATCATATCTCTTTCATCTGCTGCTATGCTATGCTATGCTATGCTATGCTATGCTATG 299
 Db 246 eIleIlePheLeuPheIleLeuLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTy 266
 QY 300 CCAAGGTATGTTGCCGTTTGTCTCTACTTCCAGGAACATCAACACACGACGCGGGCC 359
 Db 266 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrGlyPr 286
 QY 360 ATGCAAGACCTGCACGACTCTCTCTCAAGGAAACTCTAGCTTTCCCTCTTGTGTGTATC 419
 Db 286 oCysIysThrCysThrThrProAlaGlnGlyThrSerMetPheProSerCysCysTh 306
 QY 420 AAAACCTTCGACGGAACCTGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATG 479
 Db 306 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAla 326
 QY 480 ATTCTATGGAGTGGGCTCAGTCCGTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539
 Db 326 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValProPheVa 346
 QY 540 TCAGTGTGCTGAGGGCTTTCCCTCCACTGTTGGCTTTTCAGTTATGATGATGATGATGATG 599
 Db 346 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 366
 QY 600 TTGGGGGGGAGTGTGTACAACTGTGAGTCCCTTTTACCTCTATTACCAATTTCTT 659
 Db 366 rTrpGlyProSerLeuTyAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh 386
 QY 660 TTGCTCTTTGGGTATACATT 678
 Db 386 eCysLeuTrpValTyIle 392

RESULT 11

Q9W966 HPBV0
 ID Q9W966 HPBV0 PRELIMINARY; PRT; 226 AA.
 AC Q9W966;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
 DE Main protein (small surface polyprotein) (Surface antigen) (small s
 protein).
 GN Name=S;
 GN Synonyms=OG17, OG18, OG22, OG25, OG27, OG31, OG32, OG33, OG36, OG38,
 GN OG39, OG40, OG42;
 OS Hepatitis B virus.
 OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
 OC Orthohepadnavirus.
 OK NCBI_TaxID=10407;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=HBV-C5, and HBV-C4;
 RA Guo Y., Hou J.;
 RT "Establishment of the consensus sequence of hepatitis B virus
 RT prevailing in the mainland of China.";
 RL Chin. J. Microbiol. Immunol. 19:189-2000(1999).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Ogura Y., Kurosaki M., Asahina Y., Enomoto N., Marumo F., Sato C.;
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=genotype C;
 RX MEDLINE=20528425; PubMed=11074476;
 RX DOI=10.1096-9071(200012)62:4:471::AID-JMV12>3.3.CO;2-O;
 RA Rokuhara A., Tanaka E., Yagi S., Mizokami M., Hashikura Y.,
 RA Kawasaki S., Kiyosawa K.;
 RT "De novo infection of hepatitis B virus in patients with orthotopic
 RT liver transplantation: analysis by determining complete sequence of
 RT the genome.";
 RL J. Med. Virol. 62:471-478(2000).
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RA Sun X., Rokuhara A., Tanaka E., Gad A., Muto H., Matsumoto A.,
 RA Yoshizawa K., Kiyosawa K.;
 RT "Nucleotide Mutations Associated With Hepatitis B e Antigen
 RT Negativity.";
 RL J. Med. Virol. 76:170-175(2005).
 RN [5]
 RP NUCLEOTIDE SEQUENCE.
 RA Rokuhara A.;
 RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Y18856; CAB38764.1; -; Genomic DNA.
 DR EMBL; AB029967; BAB20350.1; -; Genomic DNA.
 DR EMBL; AB029971; BAB20354.1; -; Genomic DNA.
 DR EMBL; AB029974; BAB20357.1; -; Genomic DNA.
 DR EMBL; AB029976; BAB20359.1; -; Genomic DNA.
 DR EMBL; AB029980; BAB20363.1; -; Genomic DNA.
 DR EMBL; AB029981; BAB20364.1; -; Genomic DNA.
 DR EMBL; AB029982; BAB20365.1; -; Genomic DNA.
 DR EMBL; AB029985; BAB20368.1; -; Genomic DNA.
 DR EMBL; AB029987; BAB20370.1; -; Genomic DNA.
 DR EMBL; AB029988; BAB20371.1; -; Genomic DNA.
 DR EMBL; AB029989; BAB20372.1; -; Genomic DNA.
 DR EMBL; AB029991; BAB20374.1; -; Genomic DNA.
 DR EMBL; AB030507; BAA82644.1; -; Genomic DNA.
 DR EMBL; Y18855; CAB38760.1; -; Genomic DNA.
 DR EMBL; AB198079; BAD86605.1; -; Genomic DNA.
 DR EMBL; AB198080; BAD86609.1; -; Genomic DNA.
 DR EMBL; AB198081; BAD86613.1; -; Genomic DNA.
 DR EMBL; AB198082; BAD86617.1; -; Genomic DNA.
 DR EMBL; AB198083; BAD86621.1; -; Genomic DNA.
 DR EMBL; AB029966; BAB20349.1; -; Genomic DNA.
 DR PIR; JQ2094; JQ2094.
 DR PIR; JQ2095; JQ2095.
 DR PIR; JQ2096; JQ2096.
 DR PIR; JQ2097; JQ2097.

DR PIR; JQ2098; JQ2098.
 DR PIR; JQ2099; JQ2099.
 DR PIR; JQ2100; JQ2100.
 DR PIR; JQ2101; JQ2101.
 DR PIR; JQ2102; JQ2102.
 DR PIR; JQ2106; JQ2106.
 DR PIR; JQ2108; JQ2108.
 DR PIR; JQ2109; JQ2109.
 DR PIR; JQ2111; JQ2111.
 DR PIR; JQ2112; JQ2112.
 DR PIR; JQ2116; JQ2116.
 DR GO; GO:0016032; P:Viral life cycle; IEA.
 DR InterPro; IPR00349; Hepvir_surfaG.
 DR PANTHER; PTHR10832; Hepvir_surfaG; 1.
 DR Pfam; PF00695; vMSA; 1.
 KW Antigen; Polyprotein.
 SQ SEQUENCE 226 AA; 25418 MW; 2D5BEE91908A7E00 CRC64;
 Alignment Scores:
 Pred. No.: 1.63e-98 Length: 226
 Score: 1206.00 Matches: 220
 Percent Similarity: 96.92% Conservative: 0
 Best Local Similarity: 96.92% Mismatches: 6
 Query Match: 94.37% Indels: 2
 DB: 2 Gaps: 0
 US-10-761-006a-1_COPY_155_835 (1-681) x Q9W966_HPBV0 (1-226)
 QY 1 ATGGAGAACACAACTAGGATTCCTAGGACCCCTGCTGCTTACAGCGGGGTTTTC 60
 Db 1 MetGluAenThrThrSerGlyPheLeuGlyProLeuValLeuGlnAlaGlyPhePhe 20
 QY 61 TCGTTGACAAAGATCCTCACAAATACCGCAGAGTCTAGACTC-TGCTGGACTTCTCTCAAT 119
 Db 21 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpIleSerLeuAsn 40
 QY 120 TTTCTAGGGGAGCACCACCGTGTTCCTGGCCAAAATTCGCAGTCCCAACCTCCAATCA 179
 Db 41 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 60
 QY 180 CTCACCAACCTCTGTCTCCATTTGTCCTGGCTATCGCTGGAGTGTCTGCGGCGTTT 239
 Db 60 sSerProThrSerCysProProlleCysProGlyTyrArgTrpMetCysLeuArgPh 80
 QY 240 TATCATATTCTCTTCATCTGCTATGCTATCTTCTTGTGTTTCTTCGACTA 299
 Db 80 eillePheLeuPheIleLeuLeuCysleullePheLeuValLeuLeuAspTy 100
 QY 300 CCAAGGTATGTTGCCCGTTTGTCTCTACTTCCAGGAACATCAACACGACGCGGGCC 359
 Db 100 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 120
 QY 360 ATGCAAGACCTGACGACTCTCTGCTCAAGGAACTCTACGTTTCCCTCTTGTGTCTGTAC 419
 Db 120 cCysLysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysCysTh 140
 QY 420 AAAACCTTCGACGGAACATGCATTTGATTCCCATCCCATCATCTCGGCTTCGCAAG 479
 Db 140 rLysProSerAspGlyAsnCysThrCysIleProlleProSerSerTrpAlaPheAlaAr 160
 QY 480 ATTCTATGGGAGTGGGCTCAGTCCGTTTCTCTGGCTCAGTTTACTAGTGCCTATTTGT 539
 Db 160 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValProPheVa 180
 QY 540 TCAGTGGTTCGTAGGGCTTTCCCCACCTGTTTGGCTTTTCAGTTATATGATGATGTGTA 599
 Db 180 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTriPy 200
 QY 600 TTGGGGGGAAGTCTGTACAAACATCTTCAGTCCCTTTTACCTCTATTACCAATTTCTT 659
 Db 200 rTrpGlyProSerLeuTyraenilleLeuSerProPheLeuProLeuLeuPhePhe 220
 QY 660 TTGTCTTTGGGTATACATT 678

```

Db      220 eCysLeuTrpValTyrIle 226
|||||
Db      80 eIleIlePheLeuPheIleLeuLeuLeuCysLeuLeuPheLeuLeuValLeuLeuAspTy 100
|||||
QY      300 CCAAGGTATGTTGCCGTTTTCCTCTACTTCCAGGAACATCAACACGACGCGGCC 359
|||||
Db      100 rGInGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrGlyPr 120
|||||
QY      360 ATGCAGACCTGCAGACTCCCTGCCTCAAGAACTCTAGTTTCCCTCTTGTGCTGTAC 419
|||||
Db      120 oCysLysThrCysThrIleProAlaGInGlyThrSerMetPheProSerCysCysStn 140
|||||
QY      420 AAAACCTTCGACGCGAACTGCATCTGTATTCCATCCCATCATCTCTGGGCTTTCGAAG 479
|||||
Db      140 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAla 160
|||||
QY      480 ATTCTATCGGAGTGGGCTCAGTCCTGCTTTCCTGGCTCAGTTACTAGTGCATTTGT 539
|||||
Db      160 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValPropheVa 180
|||||
QY      540 TCAGTGGTTCGTAGGCTTTCCTCCCACTCTTGGCTTTCAGTTATATGATGATGTGTA 599
|||||
Db      180 lGInTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 200
|||||
QY      600 TTGGGGCGCAAGTCTGTACAACTCTTGAAGTCCCTTTTACCTCTATTACCAATTTCTT 659
|||||
Db      200 rTrpGlyProSerLeuTyrAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh 220
|||||
QY      660 TTGCTCTTGGGTATACATT 678
|||||
Db      220 eCysLeuTrpValTyrIle 226
|||||

RESULT 13
VMSA HPBVR
ID VMSA HPBVR STANDARD; PRT; 399 AA.
AC P03140;
DT 21-JUL-1986 (Rel. 01, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DE Major surface antigen precursor.
GN Name=S;
OS Hepatitis B virus (subtype adr).
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=106820;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=83168919; PubMed=6300776;
RA Ono Y., Onda H., Sasada R., Igarashi K., Sugino Y., Nishioka K.;
RT "The complete nucleotide sequences of the cloned hepatitis B virus
DNA; subtype adr and adr.";
RL Nucleic Acids Res. 11:1747-1757(1983).
RN [2]
RP NUCLEOTIDE SEQUENCE OF 175-349.
RX MEDLINE=85107103; PubMed=3968537;
RA Takeshima H., Inokoshi J., Namiki M., Shimada J., Omura S.;
RT "Structural analysis of the gene coding for hepatitis B virus surface
antigen and its product.";
J. Gen. Virol. 66:195-200(1985).
CC -1- PTM: Myristylation is essential for viral infectivity (By
similarity).
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC
CC EMBL; V00867; CAA24234.1; ALT_INIT; Genomic_DNA.
DR PIR; A03705; SAVLA.
DR InterPro; IPR000349; Hepvir_surfag.
DR PANTHER; PTHR10832; Hepvir_surfag; 1.
DR Pfam; PF00695; VMSA; 1.
DR Antigen; Glycoprotein; Lipoprotein; Myristate.
KW

Alignment Scores:
Pred. No.: 1.63e-98 Length: 226
Score: 1206.00 Matches: 220
Percent Similarity: 96.92% Conservative: 0
Best Local Similarity: 96.92% Mismatches: 6
Query Match: 94.37% Indels: 2
DB: 2 Gaps: 0

US-10-761-006a-1_COPY_155_835 (1-681) x Q9YKI9_HPBV0 (1-226)

QY 1 ATGAGAACACAAATCAGGATTCCTAGGACCCCTGCTCGTGTACAGCGGGGTTTTTC 60
Db 1 MetGluAnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20
QY 61 TCGTTGACAGAAATCTCAATACCGGAGAGTCTAGACTC-TGGTGAGCTTCTCTCAAT 119
Db 21 PheLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 40
QY 120 TTTCTAGGGGAGCACCCAGTGTCTCGGCCAAATTCGCAGTCCCAACCTCCAATCA 179
Db 41 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnH1 60
QY 180 CTCACCAACCTTGTCTCTCCCAATTTGCTCGGCTATCGCTGGATGTCTGCGGCGGTTT 239
Db 60 sSerProThrSerCysProProIleCysProGlyTyrArgTrpMetCysLeuArgGPh 80
QY 240 TATCATATTCCTTTCATCTGCTGCTATGCGTCACTCATCTTGTGTGGTCTTCTGGA 299

```

FT INIT_MET 0 0 By similarity.
FT PROPEP 1 173 Major surface antigen.
FT CHAIN 174 399 N-myristoyl glycine (by host) (By
FT LIPID 1 1 similarity)
FT CARBOHYD 14 14 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 122 122 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 176 176 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 319 319 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 399 AA; 43554 MW; 3C7437A6EEDBD4E6 CRC64;

Alignment Scores:
Pred. No.: 1.78e-98 Length: 399
Score: 1206.00 Matches: 220
Percent Similarity: 96.92% Conservative: 0
Best Local Similarity: 96.92% Mismatches: 6
Query Match: 94.37% Indels: 2
DB: 1 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x VMSA_HPBVR (1-399)

QY 1 ATGGAGAACACACATCAGGATTCCTAGGACCCCTCGTGTATACAGCGGGTTTTTC 60
DB 174 MetGluAnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 193
QY 61 TCGTTGACAAAGATCTCAATACCGCAGAGCTAGACTC-TGGTGGACTTCTCTCAAT 119
DB 194 LeuLeuThrArgileLeuThrileProGlnSerLeuAspSerTrpTrpThrSerLeuAen 213
QY 120 TTTCTAGGGGAGCACCAGCTGTTCTCGGCCAAATTCGAGTCCCGAACCTCCATCA 179
DB 214 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAenSerGlnSerProThrSerAsnHi 233
QY 180 CTCACCAACTCTTGTCCTCAATTTGCTCGCTATCGCTGAGTGTCTCGCGGTTT 239
DB 233 sSerProThrSerCysProProlleCysProGlyTyArgTrpMetCysLeuArgPh 253
QY 240 TATCATATTCTCTTCATCTGCTGCTCAATGCTCATCTTCTTGTGGTTCTTCGACTA 299
DB 253 eileilePheLeuPheileLeuLeuLeuCysLeuilePheLeuLeuValLeuLeuAspy 273
QY 300 CCAAGTATGTGCGCGTTGCTCTACTTCCAGGAACATCAACACAGCAGCGGGCC 359
DB 273 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 293
QY 360 ATGCAAGACTGCGACGACTCTGCTCAAGGAACTCTACGTTTCCCTCTTGTGTGTAC 419
DB 293 oCyslysthrCysThrileProAlaGlnGlyThrSerMetPheProSerCysCysTh 313
QY 420 AAAACCTTCGGAGCGGAACTGCACTTGTATTCATCCCATCCATCATCTCGGCTTCGCAAG 479
DB 313 rlyysProSerAspGlyAenCysThrCysileProlleProSerSerTrpAlaPheAlaAr 333
QY 480 ATTCCTATGGAGTGGGCTCAGTCCGTTTCTCTGGCTCAGTTACTAGTGCATTGT 539
DB 333 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa 353
QY 540 TCAGTGTGTGTAAGGCTTCCCGCACTGTTGGCTTTCAGTTATATGATGATGTGTA 599
DB 353 iGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 373
QY 600 TTGGGGCGGAAGTCTGACAACTTTGAGTCCCTTTTACCTCTATTACCAATTTCTT 659
DB 373 rTrpGlyProSerLeuTyAsnileLeuSerProPheLeuProLeuLeuProIlePhePh 393
QY 660 TTGCTTTTGGGTATACATT 678
DB 393 eCysLeuTrpValTyIle 399

RESULT 14
Q68RP7_HPBVO PRELIMINARY; PRT; 400 AA.
ID Q68RP7_HPBVO
AC Q68RP7;

DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Pres1/pres2 surface.
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
RN NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Liu C.-J., Chen P.-J., Chen D.-S., Kao J.-H., Lai M.-Y.;
RT "Origin of Serum Hepatitis B Virus in Acute Exacerbation: Comparison
RT with HBV in the liver and from other Exacerbation.";
RL Hepatology 40:310-317(2004).
DR EMBL; AY596108; AAU01945.1; -; Genomic DNA.
DR GO; GO:0016032; P:Viral life cycle; IEA.
DR InterPro; IPR000349; Hepvir_surfaG.
DR PANTHER; PTHR10832; Hepvir_surfaG; 1.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.
SQ SEQUENCE 400 AA; 43768 MW; DBDC9D063A569678 CRC64;

Alignment Scores:
Pred. No.: 1.78e-98 Length: 400
Score: 1206.00 Matches: 220
Percent Similarity: 96.92% Conservative: 0
Best Local Similarity: 96.92% Mismatches: 6
Query Match: 94.37% Indels: 2
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x Q68RP7_HPBVO (1-400)

QY 1 ATGGAGAACACACATCAGGATTCCTAGGACCCCTCGTGTATACAGCGGGTTTTTC 60
DB 175 MetGluAnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 194
QY 61 TCGTTGACAAAGATCTCTCAATACCGCAGAGCTAGACTC-TGGTGGACTTCTCTCAAT 119
DB 195 LeuLeuThrArgileLeuThrileProGlnSerLeuAspSerTrpTrpThrSerLeuAen 214
QY 120 TTTCTAGGGGAGCACCAGCTGTTCTCGGCCAAATTCGAGTCCCGAACCTCCATCA 179
DB 215 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAenSerGlnSerProThrSerAsnHi 234
QY 180 CTCACCAACTCTCTGCTCTCAATTTGCTCGGCTATCGCTGAGTGTCTCGCGGTTT 239
DB 234 sSerProThrSerCysProProlleCysProGlyTyArgTrpMetCysLeuArgPh 254
QY 240 TATCATATTCTCTTCATCTGCTGCTATGCTCATCTTCTTGTGGTTCTTCGACTA 299
DB 254 eileilePheLeuPheileLeuLeuLeuCysLeuilePheLeuLeuValLeuLeuAspy 274
QY 300 CCAAGTATGTGCGCGTTGCTCTACTTCCAGGAACATCAACACAGCAGCGGGCC 359
DB 274 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 294
QY 360 ATGCAAGACTGCGACGACTCTGCTCAAGGAACTCTACGTTTCCCTCTTGTGTGTAC 419
DB 294 oCyslysthrCysThrileProAlaGlnGlyThrSerMetPheProSerCysCysTh 314
QY 420 AAAACCTTCGGAGCGGAACTGCACTTGTATTCATCCCATCCATCATCTCGGCTTCGCAAG 479
DB 314 rlyysProSerAspGlyAenCysThrCysileProlleProSerSerTrpAlaPheAlaAr 334
QY 480 ATTCCTATGGAGTGGGCTCAGTCCGTTTCTCTGGCTCAGTTACTAGTGCATTGT 539
DB 334 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa 354
QY 540 TCAGTGTGTGTAAGGCTTCCCGCACTGTTGGCTTTCAGTTATATGATGATGTGTA 599
DB 354 iGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 374
QY 600 TTGGGGCGGAAGTCTGTGACAACTCTGAGTCCCTTTTACCTCTATTACCAATTTCTT 659

Db 374 rTrpGlyProSerLeuTyrAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh 394
QY 660 TTGTCTTTGGGTATACATT 678
Db 394 eCysLeuTrpValTyrIle 400

RESULT 15

Q7TDR4_HPBVO
ID Q7TDR4_HPBVO PRELIMINARY; PRT; 400 AA.
AC Q7TDR4;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PreS1/proS2 surface protein.
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Liu C.-J., Chen P.-J., Lai M.-Y., Kao J.-H., Chen D.-S.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY167095; AA041312.1; -; Genomic DNA.
DR GO; GO:0016032; P: viral life cycle; IEA.
DR InterPro; IPR000349; Hepvir_surfaG.
DR PANTHER; PTHR10832; Hepvir_surfaG; 1.
DR Pfam; PF00695; vMSA; 1.
KW Antigen.
SQ SEQUENCE 400 AA; 43694 MW; CCF173EA4365E414 CRC64;

Alignment Scores:

Pred. No.: 1.78e-98 Length: 400
Score: 1206.00 Matches: 220
Percent Similarity: 96.92% Conservative: 0
Best Local Similarity: 96.92% Mismatches: 6
Query Match: 94.37% Indels: 2
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x Q7TDR4_HPBVO (1-400)

QY 1 ATGAGACACAAATCAGGATTCCTAGGACCCCTGCTCGTTACAGGGGGTTTTC 60
Db 175 MetGluAsnThrThrSerGlyPheLeuGlyProLeuValLeuGlnAlaGlyPhePhe 194
QY 61 TCGTTGACAAATCCCTCACAAATACCGAGAGTCTAGACTC-TGGTGGACTTCTCTCAAT 119
Db 195 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 214
QY 120 TTCTAGGGGAGCACCACGCTGTTCTCGGCAAAATTCGAGTCCCAACCTCCAATCA 179
Db 215 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 234
QY 180 CTCACCAACCTCTGCTCCAAATTTGCTCGGCTATCGCTGGATGTCGTGGGGCGTTT 239
Db 234 sSerProThrSerCysProIleCysProGlyTyrArgTrpMetCysLeuArgArgPh 254
QY 240 TATCATATTCCTCTCATCTGCTGCTATGCTCATCTTCTTGTGTTCTTCTGGACTA 299
Db 254 eIleIlePheLeuPheIleLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 274
QY 300 CCAAGGTATGTGCGCGTTTGTCTCTACTTCCAGGAACATCAACACACGACGCGGCC 359
Db 274 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 294
QY 360 ATGCAAGACCTGCGACGACTCTGCTCAAGAAACTCTACGTTTCCCTCTTCTGCTGTAC 419
Db 294 oCysAlaThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysTh 314
QY 420 AARACCTTCGACGAGAACTGACTTGTATTCCTCATCCATCCATCTGGGCTTTCGCAAG 479
Db 314 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaAr 334

QY 480 ATTCTCTATGGAGTGGGCTCAGTCCGTTTCTCTGGCTCAGTTTACTAGTCCATTGCT 539
Db 334 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValProPheVa 354
QY 540 TCAGTGGTTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTTCAGTTATATGATGATGGA 599
Db 354 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 374
QY 600 TTGGGGGCGAGTCTGTACACATCTTGGTCCCTTTTACCTCTATTACCAATTTCCT 659
Db 374 rTrpGlyProSerLeuTyrAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh 394
QY 660 TTGTCTTTGGGTATACATT 678
Db 394 eCysLeuTrpValTyrIle 400

Search completed: December 27, 2005, 21:00:24

Job time : 234.454 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 27, 2005, 20:42:07 ; Search time 38.59 Seconds
(without alignments)
3395.886 Million cell updates/sec

Title: US-10-761-006A-1_COPY_155_835
Perfect score: 1278
Sequence: 1 ATGGAGACACACATCAGG.....GTCTTTGGGTATACATTAA 681

Scoring table:

BLOSUM62	Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US10761006/runat_27122005_192901_15763/app_query.fasta_1.1102
-DB=PIR -OPMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NOR=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10761006 @CNG 1 1 77 @runat_27122005_192901_15763 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DSV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR_80.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1206	94.4	400	1 SAVLA	large surface anti
2	1201	94.0	226	2 JQ2112	surface antigen -
3	1201	94.0	226	2 JQ2106	surface antigen -
4	1200	93.9	226	2 JQ2110	surface antigen -
5	1198	93.7	226	2 JQ2108	surface antigen -
6	1197	93.7	226	1 JQ1576	major surface anti
7	1197	93.7	226	1 SAVLAD	major surface anti
8	1197	93.7	226	2 JQ2104	surface antigen -
9	1197	93.7	226	2 JQ2111	surface antigen -
10	1197	93.7	226	2 JQ2098	surface antigen -
11	1196	93.6	226	1 SAVLN1	major surface anti
12	1192	93.3	226	2 JQ2115	surface antigen -
13	1191	93.2	226	2 JQ2094	surface antigen -
14	1190	93.1	226	2 JQ2101	surface antigen -

15	1190	93.1	384	2 T13469	large surface anti
16	1189	93.0	226	2 JQ2103	surface antigen -
17	1189	93.0	226	2 JQ2114	surface antigen -
18	1188	93.0	226	1 JQ1581	major surface anti
19	1188	93.0	226	2 JQ2109	surface antigen -
20	1188	93.0	226	2 JQ2105	surface antigen -
21	1187	92.9	226	2 JQ2116	surface antigen -
22	1187	92.9	226	2 JQ2100	surface antigen -
23	1186	92.8	226	2 JQ2102	surface antigen -
24	1186	92.8	400	2 S35528	surface antigen -
25	1185	92.7	226	2 JQ2099	surface antigen -
26	1184	92.6	226	1 JQ1574	major surface anti
27	1184	92.6	226	1 SAVLAR	major surface anti
28	1182	92.5	226	2 JQ2095	surface antigen -
29	1180	92.3	226	2 JQ2097	surface antigen -
30	1174	91.9	226	2 JQ2054	surface antigen -
31	1171	91.6	226	2 JQ2052	surface antigen -
32	1168	91.4	226	1 SAVLHV	major surface anti
33	1167	91.3	226	2 JQ2050	surface antigen -
34	1166	91.2	226	2 JQ2096	surface antigen -
35	1163	91.0	226	1 JQ1577	major surface anti
36	1163	91.0	226	2 JQ2053	surface antigen -
37	1163	91.0	226	2 JQ2051	surface antigen -
38	1163	91.0	400	1 JQ1575	major surface anti
39	1163	91.0	400	1 SAVLVD	large surface anti
40	1162	90.9	226	2 JQ2048	surface antigen -
41	1158	90.6	226	2 JQ2113	surface antigen -
42	1157	90.5	226	2 JQ2055	surface antigen -
43	1156	90.5	389	1 SAVLVE	large surface anti
44	1149	89.9	226	2 JQ2078	surface antigen -
45	1149	89.9	226	2 JQ2045	surface antigen -

ALIGNMENTS

RESULT 1

SAVLA

large surface antigen - hepatitis B virus (subtype adr)
N;Contains: major surface antigen; middle surface antigen

C;Species: hepatitis B virus, HBV

C;Date: 14-Nov-1983 #sequence revision 14-Nov-1983 #text_change 09-Jul-2004

C;Accession: A03705; S04569; JQ2107; PQ0608

R;Ono, Y.; Onda, H.; Sasada, R.; Igarashi, K.; Sugino, Y.; Nishioka, K.

Nucleic Acids Res. 11, 1747-1757, 1983

A;Title: The complete nucleotide sequences of the cloned hepatitis B virus DNA; subtype

A;Reference number: A93460; MUID:83168919; PMID:6300776

A;Accession: A03705

A;Molecule type: DNA

A;Residues: 1-400 <ONO>

A;Cross-references: UNIPROT:P03140; UNIPARC:UPI00001389C1; GB:V00867

R;Rho, H.M.; Kim, K.; Hyun, S.W.; Kim, Y.S.

Nucleic Acids Res. 17, 2124, 1989

A;Title: The nucleotide sequence and reading frames of a mutant hepatitis B virus subty

A;Reference number: S04568; MUID:89183619; PMID:2928116

A;Accession: S04569

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-50,'R','Y','52-66','Y','69-129','V','131-142','P','144-164','S','178-338,

A;Cross-references: UNIPARC:UPI000017496C; EMBU:X14193

R;Norder, H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnius

J. Gen. Virol. 74, 1341-1348, 1993

A;Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origi

A;Reference number: JQ2044; MUID:93329382; PMID:8336122

A;Accession: JQ2107

A;Molecule type: DNA

A;Residues: 175-400 <NOR>

A;Cross-references: UNIPARC:UPI0000033737

R;Norder, H.; Courouce, A.M.; Magnius, L.O.

J. Gen. Virol. 73, 3141-3145, 1992

A;Title: Molecular basis of hepatitis B virus serotype variations within the four major

A;Reference number: PQ0453; MUID:93107848; PMID:1469353

A;Accession: PQ0608

Db 220 eCysLeuTrpValTyrIle 226
|||||
RESULT 3
JQ2106
surface antigen - hepatitis B virus (subtype adr, strain OLL)
C;Species: hepatitis B virus, HBV
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 31-Dec-2004
C;Accession: JQ2106
R;Order, H.; Hammes, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnus,
J. Gen. Virol. 74, 1341-1348, 1993
A;Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A;Reference number: JQ2044; MUID:93329382; PMID:8336122
A;Contents: genogroup C
A;Accession: JQ2106
A;Molecule type: DNA
A;Residues: 1-226 <NOR>
A;Cross-references: UNIPROT:Q8UFP3; UNIPROT:Q91546; UNIPROT:Q9B6S4; UNIPROT:Q9DUK5; UNIP
ROT:Q9PXA0; UNIPROT:Q9DKP5; UNIPROT:Q917V3; UNIPROT:Q91557; UNIPROT:Q99H08; UNIPROT:Q9B
DH80; UNIPROT:Q8VIM3; UNIPROT:Q9DTC4; UNIPROT:Q9DKQ2; UNIPROT:Q9DUK6; UNIPROT:Q918J5; UN
C;Genetics:
A;Gene: S
C;Keywords: surface antigen

Alignment Scores:
Pred. No.: 8.1e-100 Length: 226
Score: 1201.00 Matches: 219
Percent Similarity: 96.48% Conservative: 0
Best Local Similarity: 96.48% Mismatches: 7
Query Match: 93.97% Indels: 2
DB: 2 Gaps: 0

US-10-761-006a-1_COPY_155_835 (1-681) x JQ2106 (1-226)

QY 1 ATGGAGAACACATCAGGATTCCTAGGACCCCTGCTGTTACAGGCGGGTTTTTC 60
Db 1 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20
QY 61 TCGTTGACAGAAATCCTCACAAATACCGCAGAGCTAGACTC-TGGTGGACTTCTCTCAAT 119
Db 21 TrpLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 40
QY 120 TTCTAGGGGGAGCACCACGCTGTTCTGCGCAAAATTCGACGTCCTCCCACTCCCAATCA 179
Db 41 PheLeuGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 60
QY 180 CTCACCAACTCTTGCTCCAAATTTGCTGCTATCGCTGGATGTCGCGGGTTTT 239
Db 60 sSerProThrSerCysProProIleCysProGlyTyrArgTrpMetCysLeuArgPh 80
QY 240 TATCATATTCCTTTCATCTGCTGCTATGCTCATCTTCTTCTGTTCTTCTGACTA 299
Db 80 eIleIlePheLeuPheIleLeuLeuCysLeuIlePheLeuValLeuLeuAspTy 100
QY 300 CCAAGTATGTTCCCGTTGCTCTACTTCCAGGAACATCAACACGACGCGGGCC 359
Db 100 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrGlyPr 120
QY 360 ATGCAAGACTGACGACTCTGCTCAAGGAACCTTACGTTTCCCTTCTGTTGCTGTAC 419
Db 120 oCysLysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysTh 140
QY 420 AAAACCTTCGGACGGAAACTGCTATTCATCCCATCCCATCATCTGGGCTTCGCAAG 479
Db 140 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAla 160
QY 480 ATTCCCTATGGAGTGGGCTCAGTCCGTTTCTCTGCTCAGTTACTAGTGCATTGT 539
Db 160 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValProPheVa 180
QY 540 TCAGTGGTTCGTAGGGGCTTTCCCGCTGCTGCTTTCAGTTATGAGATGATGGTA 599
Db 180 lGlnTrpPheAlaGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 200

QY 600 TTGGGGCGAGCTCTGTACACATCTTGAGTCCCTTTTACCTCTATTACCAATTTCTT 659
Db 200 rTrpGlyProSerLeuTyrAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh 220
QY 660 TTGTCTTTGGGTATACATT 678
Db 220 eCysLeuTrpValTyrIle 226

RESULT 4

JQ2110
surface antigen - hepatitis B virus (subtype adr, strain pNDR260)
C;Species: hepatitis B virus, HBV
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 31-Dec-2004
C;Accession: JQ2110
R;Order, H.; Hammes, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnus,
J. Gen. Virol. 74, 1341-1348, 1993
A;Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A;Reference number: JQ2044; MUID:93329382; PMID:8336122
A;Contents: genogroup C
A;Accession: JQ2110
A;Molecule type: DNA
A;Residues: 1-226 <NOR>
A;Cross-references: UNIPROT:Q9QMK3; UNIPARC:UPI000017841E
C;Genetics:
A;Gene: S
C;Keywords: surface antigen

Alignment Scores:
Pred. No.: 9.96e-100 Length: 226
Score: 1200.00 Matches: 218
Percent Similarity: 96.92% Conservative: 2
Best Local Similarity: 96.04% Mismatches: 6
Query Match: 93.90% Indels: 2
DB: 2 Gaps: 0

US-10-761-006a-1_COPY_155_835 (1-681) x JQ2110 (1-226)

QY 1 ATGGAGAACACATCAGGATTCCTAGGACCCCTGCTGTTACAGGCGGGTTTTTC 60
Db 1 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20
QY 61 TCGTTGACAGAAATCCTCACAAATACCGCAGAGCTAGACTC-TGGTGGACTTCTCTCAAT 119
Db 21 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 40
QY 120 TTCTAGGGGGAGCACCACGCTGTTCTGCGCAAAATTCGACGTCCTCCCACTCCCAATCA 179
Db 41 PheLeuGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 60
QY 180 CTCACCAACTCTTGCTCCAAATTTGCTGCTATCGCTGGATGTCGCGGGTTTT 239
Db 60 sSerProThrSerCysProProIleCysProGlyTyrArgTrpMetCysLeuArgPh 80
QY 240 TATCATATTCCTTTCATCTGCTGCTATGCTCATCTTCTTCTGTTCTTCTGACTA 299
Db 80 eIleIlePheLeuPheIleLeuLeuCysLeuIlePheLeuValLeuLeuAspTy 100
QY 300 CCAAGTATGTTCCCGTTGCTCTACTTCCAGGAACATCAACACGACGCGGGCC 359
Db 100 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrThrThrSerThrGlyPr 120
QY 360 ATGCAAGACTGACGACTCTGCTCAAGGAACCTTACGTTTCCCTTCTGTTGCTGTAC 419
Db 120 oCysLysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysTh 140
QY 420 AAAACCTTCGGACGGAAACTGCTATTCATCCCATCCCATCATCTGGGCTTCGCAAG 479
Db 140 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAla 160
QY 480 ATTCCCTATGGAGTGGGCTCAGTCCGTTTCTCTGCTCAGTTACTAGTGCATTGT 539
Db 160 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValProPheVa 180

```
QY 540 TCAGTGGTTCGTAGGCTTTCCCCACAGTTTGGCTTTTCAGTTATATGATGATGTGTA 599
Db |||||
Db 180 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 200
QY 600 TTGGGGCGAAGTCTGTACAACATCTTGAGTCCCTTTTACCTCTATTACCAATTTCTT 659
Db |||||
Db 200 rTrpGlyProSerLeuTyAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh 220
QY 660 TTGCTTTGGGTATACATT 678
Db |||||
Db 220 eCysLeuTrpValTyIle 226
RESULT 5
JQ2108
surface antigen - hepatitis B virus (subtype adr, strain 747/92)
C:Species: hepatitis B virus, HBV
C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 31-Dec-2004
C:Accession: JQ2108
R:Norder, H.; Hammes, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Muahahwar, I.K.; Magnus,
J. Gen. Virol. 74, 1341-1348, 1993
A:Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A:Reference number: JQ2044; MUID:93329382; PMID:8336122
A:Contents: genogroup C
A:Accession: JQ2108
A:Molecule type: DNA
A:Residues: 1-226 <NOR>
A:Cross-references: UNIPROT:Q91546; UNIPROT:Q9E84; UNIPROT:Q9DUK5; UNIPROT:Q99215; UNIP
PROT:Q9DKP5; UNIPROT:Q917V3; UNIPROT:Q91557; UNIPROT:Q99H08; UNIPROT:Q9E950; UNIPROT:Q9D
17V2; UNIPROT:Q9VM3; UNIPROT:Q9DTC4; UNIPROT:Q9DKQ2; UNIPROT:Q9DUK6; UNIPROT:Q918J5; UN
C:Genetics:
A:Gene: S
C:Keywords: surface antigen
Alignment Scores:
Pred. No.: 1.51e-99 Length: 226
Score: 1198.00 Matches: 218
Percent Similarity: 96.9% Conservative: 2
Best Local Similarity: 96.04% Mismatches: 6
Query Match: 93.74% Indels: 2
DB: 2 Gaps: 0
US-10-761-006a-1_copy_155_835 (1-681) x JQ2108 (1-226)
QY 1 ATGGAGAACACAACTCAGGATTCCTAGGACCCCTGCTGCTTACAGCGGGGTTTTTC 60
Db |||||
Db 1 MetGluSerThrThrSerGlyPheLeuGlyProLeuLeuValGlnAlaGlyPhePhe 20
QY 61 TCGTTGACAAAGATCTCACATACCGCAGAGTCTAGACTC-TGGTGGACTTCTCTCAAT 119
Db |||||
Db 21 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAen 40
QY 120 TTTCTAGGGGGAGACCCACGTGTTCTCGGCCAAAATTCCGACAGTCCCAACCTCCCAATCA 179
Db |||||
Db 41 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAenSerGlnSerProThrSerAsnHi 60
QY 180 CTCACCAACTCTGTGCTCCCAATTTGTCTCGCTATCGCTGATGTGTCTGGGGGTTT 239
Db |||||
Db 60 sSerProThrSerCysProProlleCysProGlyTyArgTrpMetCysLeuAargPh 80
QY 240 TATCATATTCTCTTCATCTCTGCTGATGCCTCATCTTCTTGTGTTCTTCTTGACTA 299
Db |||||
Db 80 ellellePheLeuPheIleLeuLeuLeuCysLeullePheLeuLeuValLeuLeuAspTy 100
QY 300 CCAAGTATGTGCCCGTTGTCTCTACTTCAGGAAACATCAACACACGACGCGGGCC 359
Db |||||
Db 100 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 120
QY 360 ATCAGAGCTGCAGACTCTCTGCTCAAGGAACTCTAGTTTCCTCTGTTGTGCTGTAC 419
Db |||||
Db 120 oCysIlysthrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysTh 140
QY 420 AAAACCTCGGAGGGAAACTGCACCTTGTATTCTCCATCCCATCATCTCGGGGTTTCCAAG 479
```

```
Db |||||
Db 140 rLysProSerAspGlyAenCysThrCysIleProIleProSerSerTrpAlaPheAlaAr 160
QY 480 ATTTCTATGGAGTGGGCTTCAGTCCCTTCTCTGCTCAGTCTTACTAGTGCCTTTGT 539
Db |||||
Db 160 gPheLeuTrpGlnTrpAlaSerValArgPheSerTrpLeuSerLeuValPropheVa 180
QY 540 TCAGTGGTTCGTAGGCTTTCCCCACAGTGTGGCTTTTCAGTTATATGATGATGTGTA 599
Db |||||
Db 180 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 200
QY 600 TTGGGGCGAAGTCTGTACAACATCTTGAGTCCCTTTTACCTCTATTACCAATTTCTT 659
Db |||||
Db 200 rTrpGlyProSerLeuTyAsnIleLeuAenProPheLeuProLeuLeuProIlePhePh 220
QY 660 TTGCTTTGGGTATACATT 678
Db |||||
Db 220 eCysLeuTrpValTyIle 226
RESULT 6
JQ1576
major surface antigen - hepatitis B virus (subtype adr, strain P8)
C:Species: hepatitis B virus, HBV
A:Note: host Homo sapiens (man)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: JQ1576; PQ0605
R:Norder, H.; Hammes, B.; Loeffdahl, S.; Courouce, A.M.; Magnus, L.O.
J. Gen. Virol. 73, 1201-1208, 1992
A:Title: Comparison of the amino acid sequences of nine different serotypes of hepatitis
A:Reference number: JQ1570; MUID:92268879; PMID:1588323
A:Accession: JQ1576
A:Molecule type: DNA
A:Residues: 1-226 <NOR>
A:Cross-references: UNIPROT:Q69599; UNIPARC:UPI00000F532E; GB:X75792; NID:g416082; PIDN
R:Norder, H.; Courouce, A.M.; Magnus, L.O.
J. Gen. Virol. 73, 3141-3145, 1992
A:Title: Molecular basis of hepatitis B virus serotype variations within the four major
A:Reference number: PQ0453; MUID:93107848; PMID:1469353
A:Accession: PQ0605
A:Molecule type: DNA
A:Residues: 101-180 <NO2>
A:Cross-references: UNIPARC:UPI0000174973
A:Experimental source: subtype adrqt, Oll
C:Genetics:
A:Gene: S
C:Superfamily: hepatitis B virus surface antigen
C:Keywords: glycoprotein; surface antigen
F:3/Binding site: carbohydrate (Asn) (covalent) #status predicted
Alignment Scores:
Pred. No.: 1.85e-99 Length: 226
Score: 1197.00 Matches: 218
Percent Similarity: 96.04% Conservative: 0
Best Local Similarity: 96.04% Mismatches: 8
Query Match: 93.66% Indels: 3
DB: 1 Gaps: 0
US-10-761-006a-1_copy_155_835 (1-681) x JQ1576 (1-226)
QY 1 ATGGAGAACACAACTCAGGATTCCTAGGACCCCTGCTGCTTACAGCGGGGTTTTTC 60
Db |||||
Db 1 MetGluAenThrThrSerGlyPheLeuGlyProLeuLeuValGlnAlaGlyPhePhe 20
QY 61 TCGTTGACAAAGATCTCACATACCGCAGAGTCTAGACTC-TGGTGGACTTCTCTCAAT 119
Db |||||
Db 21 TrpLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAen 40
QY 120 TTTCTAGGGGGAGACCCACGTGTTCTCGGCCAAAATTCCGACAGTCCCAACCTCCCAATCA 179
Db |||||
Db 41 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAenSerGlnSerProThrSerAsnHi 179
QY 120 TTTCTAGGGGGAGACCCACGTGTTCTCGGCCAAAATTCCGACAGTCCCAACCTCCCAATCA 179
Db |||||
Db 41 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAenSerGlnSerProThrSerAsnHi 60
QY 180 CTCACCAACTCTGTGCTCCCAATTTGTCTCGCTATCGCTGATGTGTCTGGGGGTTT 239
Db |||||
```

```
Db      60 sSerProThrSerCysProPheCysProGlyTyrArgTrpMetCysLeuArgArgPh 80
QY      240 TATCATATTCTCTTCATCTGCTGCTATGCTGCTCATCTTCTTGTTGGTCTCTTGACTA 299
Db      80 eileilePheLeuPheileLeuLeuLeuCysLeuilePheLeuLeuValLeuLeuAspTy 100
QY      300 CCAAGGTATGTTGCGCGTTTGTCTCTACTTCCAGGAACATCAACACCAGCAGCGGGCC 359
Db      100 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 120
QY      360 ATGCAAGACCTGCACGACTCTCTCTCAAGGAACTCTACGTTTCCCTCTTGTGTGTGTAC 419
Db      120 oCysIysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysCysTh 140
QY      420 AAAACCTTCGGACGGAACCTGACTGTATTCCTCATCCCATCATCTCTGGGCTTTCGAAG 479
Db      140 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaAr 160
QY      480 ATTCTATGGAGTGGCGCTCAGTCCGTTTCTCTCGCTCAGTTTACTAGTGCATTGCT 539
Db      160 gPheLeuTrpGluTrpAlaSerAlaArgPheSerTrpLeuSerLeuLeuValProPheVa 180
QY      540 TCAGTGTGTTGCTAGGCGCTTTCCTCCACTGTTTGGCTTTTCAGTTATATGATGATGTGTA 599
Db      180 lGlnTrpPheAlaGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 200
QY      600 TTGGGGCGGAGCTCTTACAACTCTGTAGTCCCTTTTACCTCTATTAACCAATTTCTT 659
Db      200 rTrpGlyProSerLeuTyAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh 220
QY      660 TTGCTCTTGGGTATACATT 678
Db      220 eCysLeuTrpValTyIle 226

RESULT 7
SAVLAD
major surface antigen - hepatitis B virus (subtype ad)
C:Species: hepatitis B virus, HBV
A:Note: host Homo sapiens (man)
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C:Accession: PL0053
R:Okamoto, H.; Omi, S.; Wang, Y.; Itoh, Y.; Tsuda, F.; Tanaka, T.; Akahane, Y.; Miyakawa
Mol. Immunol. 26, 197-205, 1989
A:Title: The loss of subtypic determinants in alleles, d/y or w/r, on hepatitis B surfac
A:Reference number: PL0053; MUID:89143494; PMID:2465492
A:Accession: PL0053
A:Molecule type: DNA
A:Residues: 1-226 <OKA>
A:Cross-references: UNIPROT:P31868; UNIPARC:UPI00001389BA; GB:M27765; NID:9329706; PIDN:
C:Genetics:
A:Gene: S
C:Superfamily: hepatitis B virus surface antigen
P:3/Binding site: carbohydrate (Aen) (covalent) #status predicted

Alignment Scores:
Pred. No.: 1.85e-99 Length: 226
Score: 1197.00 Matches: 218
Percent Similarity: 96.04% Conservative: 0
Best Local Similarity: 96.04% Mismatches: 8
Query Match: 93.66% Indels: 2
DB: 1 Gaps: 0

US-10-761-006a-1_COPY_155_835 (1-681) x SAVLAD (1-226)

QY      1 ATGGAACACACATCAGATTCTTAGGACCCCTGCTGTTTACAGCGGGGTTTTC 60
Db      1 MetGluAsnIleThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20
QY      61 TCCTTGACAGAACTCTCAATACCGCAGGCTAGACTC-TGGTGGACTTCTCTCAAT 119
Db      21 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 40
```

```
QY      120 TTTTAGGGGAGCACCACAGCTGTTCTCGCCAAAATTCGACAGTCCCCAACTCCCAATCA 179
Db      41 PheLeuGlyIleAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHl 60
QY      180 CTCACCAACTCTTGTCTCTCAATTTGTCTGGCTATCGCTGATGTGTCTGGCGGTTT 239
Db      60 sSerProThrSerCysProPheCysProGlyTyrArgTrpMetCysLeuArgArgPh 80
QY      240 TATCATATTCTCTTCATCTGCTGCTATGCTGCTCATCTTCTTGTTGGTCTCTTGACTA 299
Db      80 eileilePheLeuPheileLeuLeuLeuCysLeuilePheLeuLeuValLeuLeuAspTy 100
QY      300 CCAAGGTATGTTGCGCGTTTGTCTCTACTTCCAGGAACATCAACACCAGCAGCGGGCC 359
Db      100 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 120
QY      360 ATGCAAGACCTGCACGACTCTCTCTCAAGGAACTCTACGTTTCCCTCTTGTGTGTGTAC 419
Db      120 oCysIysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysCysTh 140
QY      420 AAAACCTTCGGACGGAACCTGACTGTATTCCTCATCCCATCATCTCTGGGCTTTCGAAG 479
Db      140 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaAs 160
QY      480 ATTCTATGGAGTGGCGCTCAGTCCGTTTCTCTCGCTCAGTTTACTAGTGCATTGCT 539
Db      160 nPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa 180
QY      540 TCAGTGTGTTGCTAGGCGCTTTCCTCCACTGTTTGGCTTTTCAGTTATATGATGATGTGTA 599
Db      180 lGlnTrpPheAlaGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 200
QY      600 TTGGGGCGGAGCTCTTACAACTCTGTAGTCCCTTTTACCTCTATTAACCAATTTCTT 659
Db      200 rTrpGlyProSerLeuTyAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh 220
QY      660 TTGCTCTTGGGTATACATT 678
Db      220 eCysLeuTrpValTyIle 226

RESULT 8
JQ2104
surface antigen - hepatitis B virus (subtype adr, strain pHBV1-1)
C:Species: hepatitis B virus, HBV
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 09-Jul-2004
C:Accession: JQ2104
R:Norder, H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnius
J. Gen. Virol. 74, 1341-1348, 1993
A:Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origi
A:Reference number: JQ2044; MUID:9329382; PMID:8336122
A:Contents: genogroup C
A:Accession: JQ2104
A:Molecule type: DNA
A:Residues: 1-226 <NOR>
A:Cross-references: UNIPROT:Q81100; UNIPROT:Q930V8; UNIPROT:Q81181; UNIPARC:UPI00000F3B
C:Genetics:
A:Gene: S
C:Superfamily: hepatitis B virus surface antigen
C:Keywords: surface antigen

Alignment Scores:
Pred. No.: 1.85e-99 Length: 226
Score: 1197.00 Matches: 218
Percent Similarity: 96.48% Conservative: 1
Best Local Similarity: 96.04% Mismatches: 7
Query Match: 93.66% Indels: 2
DB: 2 Gaps: 0

US-10-761-006a-1_COPY_155_835 (1-681) x JQ2104 (1-226)

QY      1 ATGGAACACACATCAGATTCTTAGGACCCCTGCTGTTTACAGCGGGGTTTTC 60
Db      1 MetGluSerThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20
```

Qy	61	TCGTGCAGAGAAATCCTCACAATACCGCAGAGCTAGACTC-TGGTGGACTTCTCTCAAT	119
Db	21	LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn	40
Qy	120	TTTTCTAGGGGGAGACCCACCGTGTTCCTGGCCAAAAATTGCGACTGCCCAACTCCCAATCA	179
Db	41	PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi	60
Qy	180	CTCACCACACTCTTGTCCTCCAAATTTGTCTCGCTATCGCTGGATGTCGGGGCTTT	239
Db	60	sSerProThrSerCysProProIleCysProGlyTyArgTrpMetCysLeuArgArgPh	80
Qy	240	TATCATATTCTCTTCACTCGCTGTGTATGCCCTCATCTTCTGTGTGTTCTTCTGGACTA	299
Db	80	elleIlePheLeuPheIleLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTy	100
Qy	300	CCAAAGTATCTGCCCGCTTTGTCTCTACTTCAGGAACAATCAACCAACGACGGGGCC	359
Db	100	xGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr	120
Qy	360	ATGCAAGACTGACGAGACTCCTGCTCAAGAAACTCTACGTTTCCCTCTGTGTCGTGTAC	419
Db	120	oCysLysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysTh	140
Qy	420	AAAACCTTCGACGGAACCTGCATCTGATTCCTCCATCCCATCATCTGGGGCTTTCGCAAG	479
Db	140	rLysProSerAspGlyAsnCysThrCysIleProlleProSerSerTrpAlaPheAlaAr	160
Qy	480	ATTCTATGGAGGPGGGCCCTCAGTCCGTTTCTCTGGCTCAGTTTACTAGTGCATTTGT	539
Db	160	gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValPropheVa	180
Qy	540	TCAGTGGTTCGTAGGGCTTTCCCCACTGTTTGGCTTTCAGTTATATGGAATGATGCGTA	599
Db	180	lGlnTrpPheAlaGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy	200
Qy	600	TTGGGGCGGAAGTCTGTACAACTCTTTGAGTCCCTTTTACCTCTATTACCAATTTTCTT	659
Db	200	rTrpGlyProSerLeuTyraIleLeuSerProPheLeuProLeuLeuProIlePhePh	220
Qy	660	TTGTCTTTGGGTATACATT	678
Db	220	eCysLeuTrpValTyIle	226

RESULT 9

JQ2111
surface antigen - hepatitis B virus (subtype adr, strain Bes)
C:Species: hepatitis B virus, HBV
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 31-Dec-2004
C:Accession: JQ2111
R:Norder, H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnus,
J. Gen. Virol. 74, 1341-1348, 1993
A:Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A:Reference number: JQ2044; MUID:95329382; PMID:8356122
A:Contents: genogroup C
A:Accession: JQ2111
A:Molecule type: DNA
A:Residues: 1-226 <NOR>
A:Cross-references: UNIPROT:Q9B6S4; UNIPROT:Q9DUK5; UNIPROT:Q99215; UNIR
PROT:Q9DKP5; UNIPROT:Q917V3; UNIPROT:Q91557; UNIPROT:Q99HU8; UNIPROT:Q9E9S0; UNIPROT:Q9D
1V2; UNIPROT:Q8VIM3; UNIPROT:Q9DTU4; UNIPROT:Q9DKQ2; UNIPROT:Q9DUK6; UNIPROT:Q918J5; UN
C:Genetics:
A:Gene: S
C:Keywords: surface antigen

US-10-761-006A-1_COPY_155_835 (1-681) x J021111 (1-226)

Qy	1	ATGAGAACACACATCAGGATTCCTAGGACCCCTGCTCGTGTTACAGGCGGGTTTTTC	60
Db	1	MetAspLysThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe	20
Qy	61	TCGTTGACAGAATCCTCACAAATACCGCAGAGTCTAGACTC-TGGTGAGACTTCTCTCAAT	119
Db	21	LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn	40
Qy	120	TTTCTAGGGGAGACCCACCGTGTCTCTGGCCAAATTCGCAGTCCCAACCTCCAATCA	179
Db	41	PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi	60
Qy	180	CTCACCAACCTTGTCTCCAAATTTGCTCGGTATCGCTGGATGTCTGCGGCGTTT	239
Db	60	sSerProThrSerCysProProIleCysProGlyTyrArgTrpMetCysLeuArgArgph	80
Qy	240	TATCATATTCCTTCATCCTCGCTGCTATGCTCATCTTCTTGTGGTTCTCTCGACTA	299
Db	80	elleIlePheLeuPheIleLeuLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTy	100
Qy	300	CCAAGGTATGTTGCCGTTTGTCTCTACTTTCAGGAACATCAACACACAGCAGGGGCC	359
Db	100	rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr	120
Qy	360	ATGCAAGACCTGCACGACTCTGCTCAAGGAAACTCTACGTTTCCCTCTTGTGTGTGTAC	419
Db	120	oCysLysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysCysTh	140
Qy	420	AAAACCTTCGGACGGAAACTGCACCTTGATTTCCATCCCATCCATCCTCGGGCTTCGCAAG	479
Db	140	rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaAr	160
Qy	480	ATTCTATGGAGTGGGCGCTCAGTCCGGTTTCTCTGGCTCAGTTTACTAGTGCCATTGT	539
Db	160	gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa	180
Qy	540	TCAGTGGTTCGTAGGGCTTCCCCACACGTGTTGGCTTTCAGTTATATCGATGATCTGTA	599
Db	180	IglNTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy	200
Qy	600	TTGGGGGCGAAGTCTGTACAAACATCTCAGTCCCTTTTAACTCTATTACCAATTTCTT	659
Db	200	rTrpGlyProSerLeuTyraNulleLeuSerProPheLeuProLeuLeuPhePhePh	220
Qy	660	TTGTCTTTGGGTATACATT	678
Db	220	eCysLeuTrpValTyrlle	226
RESULT 10			
J02098			
surface antigen - hepatitis B virus (subtype adr, strain Bau)			
C:Species: hepatitis B virus, HBV			
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 31-Dec-2004			
C:Accession: J02098			
R:Order, H.; Hammas, B.; Lee, S.D.; Bille, K.; Courouce, A.M.; Mushahwar, I.K.			
J. Gen. Virol. 74, 1341-1348, 1993			
A:Title: Genetic relatedness of hepatitis B viral strains of diverse geographi			
A:Reference number: J02044; PMID:93329382; PMID:9336122			

RESULT 10

JQ2098
 surface antigen - hepatitis B virus (subtype adr, strain Bau)
 C:Species: hepatitis B virus, HBV
 C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 31-Dec-2004
 C:Accession: JQ2098
 R:Norder, H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnus
 J. Gen. Virol. 74, 1341-1348, 1993
 A:Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origi
 A:Reference number: JQ2044; MUID:93329382; PMID:8336122
 A:Contents: genogroup C
 A:Accession: JQ2098
 A:Molecule type: DNA
 A:Residues: 1-226 <NOR>
 A:Cross-references: UNIPROT:Q8UVF3; UNIPROT:Q91546; UNIPROT:Q9E6S4; UNIPROT:Q9DUKS; UNIP
 PROT:Q9PXA0; UNIPROT:Q9DKP5; UNIPROT:Q917V3; UNIPROT:Q91557; UNIPROT:Q99HU8; UNIPROT:Q9
 DH0; UNIPROT:Q917V2; UNIPROT:Q9VIM3; UNIPROT:Q9DTC4; UNIPROT:Q9DKQ2; UNIPROT:Q918J5; U
 C:Genetics:
 C:Gene: S
 C:Keywords: surface antigen
 Alignment Scores:

Pred. No.:	1.85e-99	Length:	226
Score:	1197.00	Matches:	218
Percent Similarity:	96.48%	Conservative:	1
Best Local Similarity:	96.04%	Mismatches:	7
Query Match:	93.66%	Indels:	2
DB:	2	Gaps:	0

US-10-761-006A-1 COPY 155 835 (1-681) x JQ2098 (1-226)

Qy	1	ATGAGAACACAA	CATCAGGATTC	TACGACCCCT	CGTCTCGTGT	TTACAGCGCGGGT	TTTC	60
Db	1	MetGluAnThr	ThrSerGlyPhe	LeuGlyPro	LeuLeuVal	LeuGlnAla	GlyPhePhe	20
Qy	61	TCGTTGACAAGA	AACTCTCACA	CAATACCG	CAGAGTCTAG	ACTC-TGGTGG	ACTTCTCTCAAT	119
Db	21	LeuLeuThrA	rgIleLeuThr	IleProGln	SerLeuAsp	SerTrpTrp	ThrSerLeuAn	40
Qy	120	TTTCTAGGGG	GAGACCCAC	CGTGTTCCT	GGCCAAAAT	TCCGAGTCCC	CACCTCCATCA	179
Db	41	PheLeuGlyG	lyAlaProThr	Cys-ProGly	GlnAsnLeu	GlnSerPro	ThrSerAsnHi	60
Qy	180	CTCACCAAC	CTTGTCTC	CAATTGTCT	CGGCTATCG	CTGGATGTCT	GCGCGGTTT	239
Db	60	sSerProThr	SerCysPro	ProIleCys	ProGlyTyr	argTrpMet	CysLeuArgargh	80
Qy	240	TATCATAT	TCTCTTCAT	CTCTGCT	GTACTGCT	CACTCTCT	TGTGGTCTTCTGGACTA	299
Db	80	eileilePhe	LeupheIle	LeuLeuLeu	CysLeuIle	PheLeuLeuVal	LeuLeuAspty	100
Qy	300	CCAGGTAT	GTTGCCG	TTTCTC	CTACTTCC	CAGGAAC	CAACACACGACACGGGCGC	359
Db	100	rgInGlyMet	LeuProVal	CysPro	LeuLeuPro	GlyThrSer	ThrThrSerThrGlyPr	120
Qy	360	ATGCAAGA	CACTCTG	CTCTCA	AGAAACT	CTACGTTT	CCCTCTTGTGTCTGTAC	419
Db	120	oCysIleThr	CysThrIle	ProAla	GlnGlyThr	SerMetPhe	ProSerCysCysThr	140
Qy	420	AAAACCT	TCCGACG	GAACATG	CACTGTAT	TCCCATCC	CAATCTCTGGGCTTCCGCAAG	479
Db	140	rLysProSer	AspGly	AsnCys	ThrCysIle	ProIlePro	SerSerTrpAlaPheAlaAr	160
Qy	480	ATTCTAT	GCGAGTGG	GCTCAGT	CCGTTTCT	CTCGGCTCAG	TTTACTAGTGCCTTTGT	539
Db	160	gPheLeuTrp	GluTrpAla	SerValarg	PheSerTrp	LeuSerLeuLeuVal	ProPheVa	180
Qy	540	TCAGTGT	TCGTAGG	CTTTCCCC	CACTGT	TTTGGCTTT	CAGTTATATGATGATGTGGTA	599
Db	180	IglnTrpPhe	ValGlyLeu	SerProThr	ValTrpLeu	SerValIle	TrpMetMetTrpy	200
Qy	600	TTGGGGG	GGAAGTCT	GTATCA	CACTCT	CAGTCCCTTT	TTTACCTCTATTACCAATTTTCTT	659
Db	200	rTrpGlyPro	SerLeuThr	AsnIle	LeuAnPro	PheLeuPro	LeuLeuPhePhe	220
Qy	660	TTGTCTTT	GGGTATAC	ATT				678
Db	220	eCysLeuTrp	ValTyrIle					226

RESULT 11

SAVLN1

SAVINT
major surface antigen - hepatitis B virus (subtype adr, strain NC-1)

C;Species: hepatitis B virus, HBV

A: Note: host *Homo sapiens* (man)

C>Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 09-Jul-2004

C:Accession: JC1002

C/ACCESSION: 001002
R/Qi. Z.H.: Yan. J.: Xiong. W.J.: Cai. L.W.

XU, Z.H.; Fan, G.; Xiong, W.G.; Cai, D.W.
Chinese Biochem. J. 4, 201-209, 1988

A>Title: Determination of the nucleotide sequence and studies on the structure of hepatitis

A:Reference number: JC1002

A:Accession: JCI002

A: Accession: UCL002
A: Molecule type: DNA

A;Molecule type: DNA
A:Residues: 1-226 <OIZ>

A:Cross-ref

A;Gene: S
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: surface antigen

Alignment Scores:		
Pred. No.:	2,286-99	Length:
Score:	196.00	Matches:
Percent Similarity:	96.04%	Conservative:
Best Local Similarity:	96.04%	Mismatches:
Query Match:	93.58%	Indels:
DB:	1	Gaps:

US-10-761-006A-1 COPY 155 835 (1-681) x SAVLN1 (1-226)

QY	1	ATGAGAACACAA	CATCAGGA	TCTTAG	ACCCTTCGTG	GTTACAGCGGG	TTTTTC	60
Db								
Db	1	MetGluAenThrAlaSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe	20					
QY	61	TCGTTGACAAGA	AATCCTCACAA	TACCCGAG	GTCCTAGACTC-TGCTGGAC	TCTCTCAAT	119	
Db								
Db	21	LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAen	40					
QY	120	TTTCTAGGGG	GAGCACCCAC	GTGTCTCGGC	CAAATAATCGCAGTCCCAAC	CTCCAATCA	179	
Db								
Db	41	PheLeuGlyGlyAlaProThrCys-ProGlyGlnAenSerGlnSerProThrSerAenHi	60					
QY	180	CTCACCAAC	CTTGTCTCCCAAT	TTGTCTCGGCTATCGCTGGATGTGTCTCGCGG	TTTT	239		
Db								
Db	60	sSerProThrSerCysProProIleCysProGlyTyArgTrpMetCysLeuAenArgPh	80					
QY	240	TATCATAT	TCTCTTCAT	TCCTGCTGCTANGCTCATCTTCTTGTTGGT	TCTTCTGGACTA	299		
Db								
Db	80	eilePheLeuPheIleLeuLeuCysLeuilePheLeuLeuValLeuLeuAspTy	100					
QY	300	CCAAGTAT	GTGTGCCGTTGTCTCTACTTCCAGGA	CAATCAACACGACACGGG	GGC	359		
Db								
Db	100	rHisGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr	120					
QY	360	ATGCAAGAC	CTGCAGACTCTCTGCTCAAGAA	AACTCTACGTTTCCTCTTGTGTCTGTAC	419			
Db								
Db	120	oCysyThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysTh	140					
QY	420	AAAACTT	TCGACGGAACTGCAC	TGTGTATTTCCCATCCCATCATCTCGGGCTT	TCGCAAG	479		
Db								
Db	140	rLysProSerAspGlyAenCysThrCysileProIleProSerSerTrpAlaPheAlaAr	160					
QY	480	ATTCTAT	GGGAGTGGGCTCAGTCCGTTTCTCTGGCTCAGTTTACTAGTGC	CAATTTGT	539			
Db								
Db	160	pHeLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa	180					
QY	540	TCAGTGT	TCTTAGGGCTTTCCCCAC	GTTTGGCTTTTCAGTTATATGATCATGTGGTA	599			
Db								
Db	180	IgInTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpy	200					
QY	600	TTGGGGG	CGAAGTCTGTACAA	CATCTTGAGTCCCTTTTACCTCTATTA	ACCAATTTCTT	659		
Db								
Db	200	rTrpGlyProSerLeuTyAenIleLeuSerProPheLeuProLeuLeuProIlePhePh	220					
QY	660	TTGCTTT	GGGTATACATT	678				
Db								
Db	220	eCysLeuTrpValTyVrille	226					

RESULT 12

RESUL
JO2115

surface antigen - hepatitis B virus (subtype adw, strain PIWK146)

C:Species: hepatitis B virus. HBV
virace antigen - hepatitis B VII

```
C;Species: непаристія в вітків, НВВ
C;Date: 03-Feb-1994 #sequence revision 03-Feb-1994 #text change 31-Dec-2004
```

C:Date: 03-FEB-1994
C:Accession: J02115

C; Accession: 002113
R: Norder, H.: Hammag. B.: Lee. S.D.

J. Gen. Virol. 74: 1341-1348. 1993.

A: Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
U. Gen. VIOL. 74, 1341-1346, 1993

A; title: Genetic relationship of hepatitis B virus strains
A:Reference number: JQ2044; MUID:93329382; PMID:83361222

A;Contents: genoqgroup C

420 A A A A C C T T C G G A C G G A A A C T G C A C T T G T A T T C C C A T C C C A T C A T C T C G G G C T T T C G C A A G 479

140 r l y s P r o S e r A s p G l y A s n C y s T h r C y l l e P r o l e P r o S e r S e r T r p A l a P h e A l a A r 160

480 A T T C C T A T G G G A C T G G C C T C A G T C C G T T T C T C T G G C T C A G T T T A C T A G T G C C A T T T G T 539

160 g P h e L e u T r p G l u T r P a l a S e r A l a A r g P h e S e r T r p L e u S e r L e u L e u V a l P r o P h e V a 180

540 T C A G T G T T C G T A G G G C T T T C C C C A C T G T T T G G C T T T C A G T A T A T G G A T G A T G T G T A 599

180 l G l n T r p P h e V a l G l y L e u S e r P r o T h r V a l T r p L e u S e r V a l l e T r p M e t T r p T y 200

600 T T C G G G C G C A A G T C T G T A C A C A C T T T G A G T C C C T T T T T A C C T C T A T T A C C A A T T T T C T T 659

200 r T r p G l y P r o S e r L e u H i s A s n L e u A s n P r o P h e L e u P r o L e u L e u P r o l i e P h e P h 220

660 T T C T C T T T G G G T A T A C A T T 678

220 e C y s L e u T r p V a l T y r I l e 226

RESULT 15

T13469

N;Alternate names: envelope protein

N;Contains: major surface antigen; middle surface antigen

C;Species: hepatitis B virus, HBV

A;Variety: isolate 09D09HCC

C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004

C;Accession: T13469

R;Takahashi, K.; Akahane, Y.; Hino, K.; Ohta, Y.; Mishiro, S.

Arch. Virol. 143, 2313-2326, 1998

A;Title: Hepatitis B virus genomic sequence in the circulation of hepatocellular carcinoma patients

A;Reference number: Z17684; MUID:99129050; PMID:9930189

A;Accession: T13469

```

A;Status: pretranslational; translated from Q97EMDZ/0500
A;Molecule type: DNA
A;Residues: 1-384 <TAK>
A;Cross-references: UNIPROT:Q9Y2U2; UNIPARC:UPI00000F1C85; EMBL:AB014368; NID:
A;Experimental source: Japanese patient with hepatocellular carcinoma isolate
C;Genetics:
A;Gene: S
A;Introns: 123/2
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: glycoprotein; surface antigen

Alignment Scores:
Pred. No.: 7.75e-99 Length: 384
Score: 1190.00 Matches: 217
Percent Similarity: 95.59% Conservatives: 0
Best Local Similarity: 95.59% Mismatches: 9
Query Match: 93.11% Indels: 2
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x T13469 (1-384)

Qy 1 ATGGAGAACACAAATCAGGATTCCTAGGACCCTCGCTCGTGTACAGGCGGGGTTTTTC 60
Db 159 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnaGlyPhePhe 178

Qy 61 TCGTTGACAAAGAATCCTCACAAATACCGCAGAGTCTAGACTC-TGGTGGACTTCTCTCAAT 119
Db 179 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTyrTyrThrSerLeuAsn 198

Qy 120 TTTCTAGGGGAGACCCACGATGTCTTCTGGCCAAAATTCGCAGTCCCCAACCTCCAATCA 179
Db 199 PheLeuGlyGlyAlaPro**Cys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 218

Qy 180 CTCACCAACCTCTTGCTCTCCAAATTTGTCCTGGCTATCGCTGGATGTGTCGCGCGGTTT 239
Db 218 sSerProThrSerCysProProIleCysProGlyTyrArgTyrMetCysLeuArgPhe 238

Qy 240 TATCATATTCCTCTTCATCTCTGCTGCTATGCCTCATCTTCTTGTGGTTCTTCTCGGACTA 299

```

```

Alignment Scores:
Pred. No.:      7.88e-99      Length:      226
Score:          1190.00      Matches:     216
Percent Similarity: 96.48%      Conservative: 3
Best Local Similarity: 95.15%      Mismatches: 7
Query Match:      93.11%      Indels:     2
DB:               2           Gaps:        0

US-10-761-006A-1_COPY_155_835 (1-681) x JQ2101 (1-226)

Qy      1  ATGAGAACACCAACATCAGGATTCCTAGGACCCCTGCTCGTTGTACAGGCGGGTTTTTC 60
Db      1  MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20
Qy      61  TCGTTGACAGAATCCTCACAAATACCGCAGAGTCAGACTC-TGGTGGACTTCTCTCAAT 119
Db      21  TrpLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 40
Qy      120  TTTCTAGGGGGAGCACCACCGTGTCTCTGGCCAAAATTTCGAGGTCCCAACCTCCCAATCA 179
Db      41  PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnLeuGlnSerProThrSerAsnHi 60
Qy      180  CTCACCAACCTCTTGTCCTCAAAATTTGTCTCGGTGATCGCTGGAGATGTCTGGCGCGTTT 239
Db      60  sSerProThrSerCysProPheCysProGlyTyzArgTrpMetCysLeuArgArgPh 80
Qy      240  TATCATATTCTCTTTCATCCGTCGTATGCCTCATCTCTGTGGTCTCTTCGGACTA 299
Db      80  eilePheLeuPheIleLeuLeuLeuCysLeuilePheLeuLeuValLeuLeuAspTy 100
Qy      300  CCAAGGTATGTTGCCCGGTTTCTCCTCTACTTTCAGGGAACATCAACACCAAGCAGCGGGCC 359
Db      100  xGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 120
Qy      360  ATCAAGACCTGCACGACTCTGCTCAAGGAAACTCTACGTTTCCCTCTTGTCGTATC 419
Db      120  ocVeaGThrCysThrThrProAlaGlnGlyThrSerMetPheProSerCysCysGlyTh 140

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 27, 2005, 20:40:32 ; Search time 18.262 Seconds
(without alignments)
3320.243 Million cell updates/sec

Title: US-10-761-006a-1_COPY_527_595

Perfect score: 139

Sequence: 1 ACACCTCTGCTCAAGAAA.....AACCTTCGACGGAACACTGC 69

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Rgapop 6.0, Rgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US10761006/runat_27122005_192900_15743/app_query.fasta_1.1102
-DB=A_Geneseq -QFT=fastan -SUPPLX=rag -MINMATCH=0.1 -LOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=BLOSUM62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USBR=US10761006@cgn_1_1_476 @runat_27122005_192900_15743 -NCPU=3
-NO MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq 21.*

1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001s.*
5: Geneseq2002s.*
6: Geneseq2003as.*
7: Geneseq2003bs.*
8: Geneseq2004s.*
9: Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	139	100.0	400	AAV54045	AAV54045 Amino aci
2	133	95.7	38	AAE17024	AAE17024 Hepatitis
3	133	95.7	48	AAE17025	AAE17025 Hepatitis
4	130	93.5	24	AAAR50965	AAAR50965 Peptide c
5	130	93.5	24	AAAR50961	AAAR50961 Amino aci
6	130	93.5	24	AAAR50968	AAAR50968 Peptide c
7	130	93.5	24	AAAR94433	AAAR94433 HBV surfa
8	130	93.5	24	AAAY04027	AAAY04027 HBsAg pep
9	130	93.5	24	AAAE25520	AAAE25520 Hepatitis

10	130	93.5	43	2	AAAR11409	AAAR11409 Hepatitis
11	130	93.5	51	9	ADW68797	ADW68797 Hepatitis
12	130	93.5	55	9	ADW68794	ADW68794 Hepatitis
13	130	93.5	60	9	ADW68793	ADW68793 Hepatitis
14	130	93.5	70	9	ADW68806	ADW68806 Hepatitis
15	130	93.5	70	9	ADW68821	ADW68821 Hepatitis
16	130	93.5	70	9	ADW68856	ADW68856 Hepatitis
17	130	93.5	71	2	AAW65505	AAW65505 Hepatitis
18	130	93.5	71	3	AAAY85118	AAAY85118 Hepatitis
19	130	93.5	75	9	ADW68792	ADW68792 Hepatitis
20	130	93.5	90	9	ADW68815	ADW68815 Hepatitis
21	130	93.5	90	9	ADW68814	ADW68814 Hepatitis
22	130	93.5	154	9	ADW68791	ADW68791 Hepatitis
23	130	93.5	154	9	ADW68813	ADW68813 Hepatitis
24	130	93.5	160	7	ADG76930	ADG76930 Hepatitis
25	130	93.5	165	7	ADG76927	ADG76927 Hepatitis
26	130	93.5	209	2	AAAR5288	AAAR5288 Deduced s
27	130	93.5	226	1	AAAP30096	AAAP30096 Sequence
28	130	93.5	226	1	AAAP30060	AAAP30060 Surface a
29	130	93.5	226	1	AAAP60223	AAAP60223 Sequence o
30	130	93.5	226	1	AAAP80417	AAAP80417 Sequence
31	130	93.5	226	2	AAAR33252	AAAR33252 HBsAg enc
32	130	93.5	226	2	AAAR55282	AAAR55282 Deduced s
33	130	93.5	226	3	AAAB35766	AAAB35766 Peptidyl
34	130	93.5	226	4	AAE04711	AAE04711 Hepatitis
35	130	93.5	226	4	AAAB98041	AAAB98041 Hepatitis
36	130	93.5	226	4	AAAB80956	AAAB80956 Viral pro
37	130	93.5	226	5	AAAM52694	AAAM52694 Hepatitis
38	130	93.5	226	5	AAAB05219	AAAB05219 Hepatitis
39	130	93.5	226	5	AAAU97646	AAAU97646 Hepatitis
40	130	93.5	226	8	ADU69200	ADU69200 Hepatitis
41	130	93.5	226	9	ADW68808	ADW68808 Hepatitis
42	130	93.5	226	9	ADY62464	ADY62464 Hepatitis
43	130	93.5	226	9	AAEA08824	AAEA08824 Hepatitis
44	130	93.5	250	2	AAAR11496	AAAR11496 RFI42/HBS
45	130	93.5	251	2	AAAR11495	AAAR11495 RFI35/HBS

ALIGNMENTS

RESULT 1

AAV54045

ID AAV54045 standard; protein; 400 AA.

XX AAV54045;

XX AC

DT 27-MAR-2000 (first entry)

XX

DE Amino acid sequence of a HBV large surface antigen protein.

XX HBV; HBV surface antigen-'S'-133 Oon strain (Met to Thr); DNA polymerase;
large surface antigen; core protein; transactivating X protein;
hepatitis vaccine; HBV infection; hepatocellular carcinoma.

XX Hepatitis B virus.

XX

FH Key Location/Qualifiers

FT Misc-difference 14

FT /note= "encoded by AGC"

FT

FT Misc-difference 174, 400

FT /note= "these residues are specifically claimed in claim

29"

FT Misc-difference 298, .320

FT /note= "these residues are specifically claimed in claim

23"

FT

XX WO9966048-A1.

XX

XX 23-DEC-1999.

XX

XX 19-JUN-1998; 98WO-SG000046.

XX

PR 19-JUN-1998; 98WO-SG000046.

XX (GOVE-) GOVERNMENT REPUBLIC SINGAPORE.
 PA Oon CJ, Lim GK, Zhao Y, Chen WN;
 PI WPI; 2000-106104/09.
 XX DR N-PSDB; AA237088.
 XX PT New isolated hepatitis B virus strain, useful for, e.g. treatment of
 PT hepatitis infection.
 XX PS Claim 23; Page 39-40; 68pp; English.
 XX The present sequence is encoded by the genome of an isolated strain of
 CC Hepatitis B virus designated human Hepatitis B virus (HBV) surface
 CC antigen-'S'-133 On strain (Met to Thr). The viral genome is deposited as
 CC ECCC accession numbers P97121501, P97121502 and P97121503. The nucleotide
 CC sequence was isolated from hepatocellular carcinoma (HCC). The nucleotide
 CC sequence encodes four overlapping proteins, which are a DNA polymerase, a
 CC large surface antigen, a core protein, and a transactivating X protein.
 CC The large surface antigen differs from the wild type sequence in that it
 CC contains a Thr at position 133 of the wild type sequence instead of a
 CC Met. The proteins are used to produce antibodies. The proteins,
 CC polynucleotide and antibodies can be used for detecting the novel HBV
 CC strain. The HBV polypeptides can also be used in hepatitis vaccines. The
 CC HBV novel strain polypeptides can be used to identify compounds for
 CC treating or preventing HBV infection or hepatocellular carcinoma
 XX Sequence 400 AA;
 SQ

Alignment Scores:
 Pred. No.: 6,24e-12 Length: 400
 Score: 139.00 Matches: 23
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 3 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x AAY54045 (1-400)

QY 1 ACGACTCTGCTCAAGAACTCTACGTTCCCTCTGTGTCTGTACAAACCTTCGGAC 60
 Db 299 ThrThrProAlaGlnGlyAsnSerThrPheProSerCysCysThrLysProSerAsp 318
 QY 61 GGAAGACTGC 69
 Db 319 GlyAsnCys 321

RESULT 2
 AAEL17024
 ID AAEL17024 standard; peptide; 38 AA.
 AC AAEL17024;
 XX 18-APR-2002 (first entry)
 DT Hepatitis B virus (HBV) adw subtype S antigen #1.

XX Hepatitis B virus; HBV; core antigen; HBcAg; immune system; typhoid;
 KW prophylactic; gene therapy; vaccine; hepatitis A virus; HAV; herpes;
 KW hepatitis C virus; HCV; influenza; foot-and-mouth disease; diarrhoea;
 KW tuberculosis; polio; rabies; acquired immunodeficiency syndrome; AIDS;
 KW dengue fever; yellow fever; malaria; whooping cough; salmonellosis;
 KW food poisoning; meningitis; gonorrhea; antiviral; antibacterial;
 KW antiprotozoal; S antigen.

OS Hepatitis B virus.

XX WO200198333-A2.

PN 27-DEC-2001.

XX 22-JUN-2001; 2001WO-GB002817.

XX 22-JUN-2000; 2000GB-00015308.
 PR 06-OCT-2000; 2000GB-00024544.
 XX (CELL-) CELLTech PHARM LTD.
 PA Page M, Li J, Pumpens P;
 XX WPI; 2002-098223/13.
 XX DR N-PSDB; AA27425.
 XX New proteins comprising a modified hepatitis B core antigen, useful as a
 PT vaccine in prophylactic or therapeutic vaccination of the human or animal
 PT body, particularly against hepatitis B virus infection.
 XX Disclosure; Fig 3B; 40pp; English.

XX The invention relates to modified proteins comprising hepatitis B virus
 CC (HBV) core antigen (HBcAg) wherein one or more of the four arginine
 CC repeats has been deleted and the protein comprising the C-terminal
 CC cysteine of HBcAg. The deleted region may be replaced by an epitope from
 CC a protein other than HBcAg, in which case the HBcAg acts as a carrier to
 CC present the epitope to the immune system. This chimeric protein or its
 CC nucleic acid is useful as a vaccine or in a method of prophylactic or
 CC therapeutic vaccination of the human or animal body, particularly against
 CC HBV. The nucleic acid encoding the protein may be used in gene therapy or
 CC DNA vaccination protocols. The chimeric protein or its nucleic acid may
 CC also be used as the basis of a prophylactic vaccine against a range of
 CC diseases, e.g. HBV, hepatitis A virus (HAV), hepatitis C virus (HCV),
 CC influenza, foot-and-mouth disease, polio, herpes, rabies, acquired
 CC immunodeficiency syndrome (AIDS), dengue fever, yellow fever, malaria,
 CC tuberculosis, whooping cough, salmonellosis, typhoid, food poisoning,
 CC diarrhoea, meningitis or gonorrhea. The present sequence is Hepatitis B
 CC virus (HBV) adw subtype S antigen

XX Sequence 38 AA;
 SQ

Alignment Scores:

Pred. No.: 4,27e-11 Length: 38
 Score: 133.00 Matches: 22
 Percent Similarity: 95.65% Conservative: 0
 Best Local Similarity: 95.65% Mismatches: 1
 Query Match: 95.68% Indels: 0
 DB: 5 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x AAEL17024 (1-38)

QY 1 ACGACTCTGCTCAAGAACTCTACGTTCCCTCTGTGTCTGTACAAACCTTCGGAC 60
 Db 16 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProSerAsp 35
 QY 61 GGAAGACTGC 69
 Db 36 GlyAsnCys 38

RESULT 3

AAEL17025
 ID AAEL17025 standard; peptide; 48 AA.

XX AAEL17025;
 AC 18-APR-2002 (first entry)

DT Hepatitis B virus (HBV) adw subtype S antigen #2.

XX Hepatitis B virus; HBV; core antigen; HBcAg; immune system; typhoid;
 KW prophylactic; gene therapy; vaccine; hepatitis A virus; HAV; herpes;
 KW hepatitis C virus; HCV; influenza; foot-and-mouth disease; diarrhoea;
 KW tuberculosis; polio; rabies; acquired immunodeficiency syndrome; AIDS;
 KW dengue fever; yellow fever; malaria; whooping cough; salmonellosis;
 KW food poisoning; meningitis; gonorrhea; antiviral; antibacterial;
 KW antiprotozoal; S antigen.

XX

OS Hepatitis B virus.
 XX PN WO200198333-A2.
 XX PD 27-DEC-2001.
 XX PF 22-JUN-2001; 2001WO-GB002817.
 XX PR 22-JUN-2000; 2000GB-00015308.
 XX PR 06-OCT-2000; 2000GB-00024544.
 XX PA (CELL-) CELLTech PHARM LTD.
 XX PI Page M, Li J, Pumpens P;
 XX DR WPI; 2002-098223/13.
 XX DR N-PSDB; AAD27426.
 XX PT New proteins comprising a modified hepatitis B core antigen, useful as a
 PT vaccine in prophylactic or therapeutic vaccination of the human or animal
 PT body, particularly against hepatitis B virus infection.
 XX PS Disclosure; Fig 3C; 40pp; English.
 XX CC The invention relates to modified proteins comprising hepatitis B virus
 CC (HBV) core antigen (HBcAg) wherein one or more of the four arginine
 CC repeats has been deleted and the protein comprising the C-terminal
 CC cysteine of HBcAg. The deleted region may be replaced by an epitope from
 CC a protein other than HBcAg, in which case the HBcAg acts as a carrier to
 CC present the epitope to the immune system. This chimeric protein or its
 CC nucleic acid is useful as a vaccine or in a method of prophylactic or
 CC therapeutic vaccination of the human or animal body, particularly against
 CC HBV. The nucleic acid encoding the protein may be used in gene therapy or
 CC DNA vaccination protocols. The chimeric protein or its nucleic acid may
 CC also be used as the basis of a prophylactic vaccine against a range of
 CC diseases, e.g. HBV, hepatitis A virus (HAV), hepatitis C virus (HCV),
 CC influenza, foot-and-mouth disease, polio, herpes, rabies, acquired
 CC immunodeficiency syndrome (AIDS), dengue fever, yellow fever, malaria,
 CC tuberculosis, whooping cough, salmonellosis, typhoid, food poisoning,
 CC diarrhoea, meningitis or gonorrhoea. The present sequence is Hepatitis B
 CC virus (HBV) adw subtype S antigen
 XX SQ Sequence 48 AA;
 Alignment Scores:
 Pred. No.: 4.33e-11 Length: 48
 Score: 133.00 Matches: 22
 Percent Similarity: 95.65% Conservative: 0
 Best Local Similarity: 95.65% Mismatches: 1
 Query Match: 95.68% Indels: 0
 DB: 5 Gaps: 0
 US-10-761-006A-1_COPY_527_595 (1-69) x AAE17025 (1-48)
 QY 1 ACGACTCTGCTCAAGGAACCTCTACGTTTCCCTCTGTGCTGTACAAACCTTCGGAC 60
 Db 16 ThrThrProAlaGlnGlyAenSerMetPheProSerCysCysThrLysProSerAsp 35
 QY 61 GGAACCTGC 69
 Db 36 GlyAenCys 38
 RESULT 4
 AAR50965
 ID AAR50965 standard; peptide; 24 AA.
 XX AC AAR50965;
 XX DT 25-MAR-2003 (revised)
 XX DT 14-OCT-1994 (first entry)
 XX DE Peptide cross reactive with anti-HBsAg antiserum.

KM Hepatitis B virus; HBV; epitope; antigen; HBsAg; diagnosis; detection;
 KM vaccine; treatment; antiserum; adw; surface antigen; viral infection.
 XX OS Synthetic.
 XX FH Key Location/Qualifiers
 FT Modified-site 10
 FT /label= Other.
 FT /note= "Met(O)."
 XX PN WO9405698-A1.
 XX PD 27-MAR-1994.
 XX PF 30-AUG-1993; 93WO-EP002342.
 XX PR 01-SEP-1992; 92AT-00001746.
 XX PA (UNNA-) UNITED NATIONS IND DEV ORG.
 XX PI Manivel V, Rao KV, Panda SK;
 XX DR WPI; 1994-101123/12.
 XX PT Peptide(s) for use in diagnosis or vaccines for hepatitis B virus -
 PT comprising oligomers of an epitope sequence of hepatitis B surface
 PT antigen.
 XX PS Claim 2; Page 14; 26pp; English.
 XX CC The peptide shows cross reactivity with anti-hepatitis B surface antigen
 CC (HBsAg) antiserum. The peptide comprises the "a" epitope of HBsAg, a
 CC group specific determinant of HBsAg. It can be used for the diagnosis of
 CC hepatitis B virus infection and in vaccines against the virus,
 CC particularly in combination with alum as adjuvant and with a myristic
 CC acid residue added to the amino terminus of the peptide. This sequence is
 CC an analogue of the sequence described in AAR50961. See also in AAR50962-
 CC 70. (Updated on 25-MAR-2003 to correct PN field.)
 XX SQ Sequence 24 AA;
 Alignment Scores:
 Pred. No.: 1.17e-10 Length: 24
 Score: 130.00 Matches: 21
 Percent Similarity: 95.65% Conservative: 1
 Best Local Similarity: 91.30% Mismatches: 1
 Query Match: 93.53% Indels: 0
 DB: 2 Gaps: 0
 US-10-761-006A-1_COPY_527_595 (1-69) x AAR50965 (1-24)
 QY 1 ACGACTCTGCTCAAGGAACCTCTACGTTTCCCTCTGTGCTGTACAAACCTTCGGAC 60
 Db 2 ThrThrProAlaGlnGlyAenSerMetPheProSerCysCysThrLysProThrAsp 21
 QY 61 GGAACCTGC 69
 Db 22 GlyAenCys 24
 RESULT 5
 AAR50961
 ID AAR50961 standard; peptide; 24 AA.
 XX AC AAR50961;
 XX DT 25-MAR-2003 (revised)
 XX DT 14-OCT-1994 (first entry)
 XX DE Amino acids 124-147 of Hepatitis B surface antigen subtype adw.
 XX KM Hepatitis B virus; HBV; epitope; antigen; HBsAg; diagnosis; detection;
 XX KW vaccine; treatment; antiserum; adw; surface antigen; viral infection.

PI Baralle PE, Scodeller E, Tisminetzky S;
 XX WPI; 1996-139691/14.
 XX
 XX New molecular presentation system - comprising a viral protein from a
 PT small insect virus in which heterologous amino acid sequences are
 PT inserted.
 XX
 XX Disclosure; Page 7; 39pp; English.
 XX
 XX The B-cell epitope (AAR94433) of the hepatitis B virus surface protein
 CC may be inserted into the outward-directed L1, L2, L3, L2 or L3 loop of
 CC the Flock House virus capsid protein (AAR88755). The capsid protein
 CC provides a conformationally suitable location for this (or other, see
 CC AAR94430-32 and AAR94434-45) heterologous peptides. A molecular
 CC presentation system is obtd. The FHV recombinant capsomer can be
 CC expressed in E. coli. Alternatively, expression in insect cells using a
 CC baculovirus vector results in prodn. of mature virus-like particles.
 CC (Updated on 27-AUG-2003 to correct OS field.)
 XX
 XX Sequence 24 AA;

Alignment Scores:
 Pred. No.: 1.17e-10 Length: 24
 Score: 130.00 Matches: 21
 Percent Similarity: 95.65% Conservative: 1
 Best Local Similarity: 91.30% Mismatches: 1
 Query Match: 93.53% Indels: 0
 DB: 2 Gaps: 0

US-10-761-006a-1_COPY_527_595 (1-69) x AAR94433 (1-24)

QY 1 ACAGACTCTCTCAAGAAACTCTACGTTTCCCTCTGTGTGTACAAACCTTCGGAC 60
 Db 2 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrllysProThrAsp 21
 QY 61 GGAAACTGC 69
 Db 22 GlyAsnCys 24

RESULT 8
 AAY04027
 ID AAY04027 standard; peptide; 24 AA.
 AC AAY04027;

17-OCT-2003 (revised)
 20-DEC-1999 (first entry)
 HBsAg peptide 124-147 for use in immunogenic complex.

Macroglobulin; immunogenic complex; hepatitis; surface antigen; HBsAg;
 antigen; vaccine.

Hepatitis B virus.

W09950303-A2

07-OCT-1999.

01-APR-1999; 99WO-US007236.

01-APR-1998; 98US-00053301.

31-MAR-1999; 99US-00282826.

(UYDU-) UNIV DUKE.

Pizzo SV, Gron H;

WPI; 1999-601324/51.

Immune response modulator alpha-2 macroglobulin complex useful in the
 treatment of HIV or hepatitis.

Claim 3; Page 77; 103pp; English.

A stable complex comprising at least one intact biomolecule and activated
 alpha-2-macroglobulin (alpha-2M) having an intact bait region is claimed,
 in which the biomolecule is covalently bound to an amino acid residue of
 a cleaved thiol ester of the alpha-2M, the amino acid residue being a
 glutamyl residue and/or a cysteinyl residue. Preparation of the complex
 comprises activation of alpha-2M with a nucleophilic compound followed by
 incubation with the intact biomolecule at elevated temperature. Covalent
 incorporation of the intact biomolecule is thus effected without use of a
 protease. The obtained complex may be used as an antigen for stimulation of
 immune response, e.g. in the form of a vaccine. Enhanced presentation of
 the biomolecule is provided, especially for those antigens which are
 poorly immunogenic. Reduction of immunodominance of particular epitopes
 is also provided. Preferably the biomolecule is an HIV antigen or a
 hepatitis virus antigen, or a peptide, fragment, hybrid or chimeric
 peptide thereof. The present sequence is a preferred example of a
 hepatitis peptide which may be used in the complex. It represents
 residues 124-147 of Hepatitis B virus surface antigen (HBsAg) (Updated on
 17-OCT-2003 to standardise OS field)

Sequence 24 AA;

Alignment Scores:
 Pred. No.: 1.17e-10 Length: 24
 Score: 130.00 Matches: 21
 Percent Similarity: 95.65% Conservative: 1
 Best Local Similarity: 91.30% Mismatches: 1
 Query Match: 93.53% Indels: 0
 DB: 2 Gaps: 0

US-10-761-006a-1_COPY_527_595 (1-69) x AAY04027 (1-24)

QY 1 ACAGACTCTCTCAAGAAACTCTACGTTTCCCTCTGTGTGTACAAACCTTCGGAC 60
 Db 2 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrllysProThrAsp 21
 QY 61 GGAAACTGC 69
 Db 22 GlyAsnCys 24

RESULT 9
 AAE25520
 ID AAE25520 standard; peptide; 24 AA.
 AC AAE25520;

06-AUG-2003 (revised)
 04-NOV-2002 (first entry)

Hepatitis b virus surface antigen (HBsAg) peptide.

Stable complex; alpha-2-macroglobulin; immunocompetence; vaccine;
 immune response; macrophage; tumour; human immunodeficiency virus; HIV;
 cytostatic; antimicrobial; alpha2M; anti-HIV; virucide; HBsAg.

Hepatitis B virus.

US6403032-B1

11-JUN-2002.

31-MAR-1999; 99US-00282826.

01-APR-1998; 98US-00053301.

(UYDU-) UNIV DUKE.

Pizzo S, Gron H;

WPI; 1999-601324/51.

PT Immune response modulator alpha-2 macroglobulin complex useful in the
PT treatment of HIV or hepatitis.
XX
PS
XX Claim 3; Col 35; 34pp; English.
XX
CC The invention relates to a stable complex comprising an intact
CC biomolecule with a nucleophilic group which is covalently bound to
CC activated alpha-2-macroglobulin (alpha2M) with an intact bait region. The
CC stable complex is useful for enhancing host immunocompetence and for
CC preparing vaccines for preventing and treating disease states. It is
CC useful for increasing immune response to a poorly immunogenic antigen,
CC and for suppressing immune response to a particular antigen. It is most
CC preferably useful for administering antigens recognised by the macrophage
CC in view of the existence on the macrophage of receptors for alpha2M. It
CC is useful for treating cancer or infection, where the isolated protein is
CC derived from tumour, abnormal cells or infectious organism, and this
CC protein can be used as an antigen and prepared in alpha2M complex which
CC is then administered in the form of vaccine to immunise mammalian
CC patients in need of such treatment. It can also be used for inducing
CC immune response against human immunodeficiency virus (HIV) antigens, and
CC hepatitis virus antigens. The present sequence is Hepatitis b virus
CC surface antigen (HBSAg) peptide used in the exemplification of the
CC invention. (Updated on 06-AUG-2003 to correct OS field.)
XX
SQ Sequence 24 AA;

Alignment Scores:
Pred. No.: 1.17e-10 Length: 24
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x AAE25520 (1-24)

QY 1 ACGACTCTGCTCAAGAACTCTACGTTTCCTCTTGTGCTGTACAAAACCTTCGGAC 60
Db 2 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProThrAsp 21

QY 61 GGAAACTGC 69
Db 22 GlyAsnCys 24

RESULT 10
AAR11409
ID AAR11409 standard; protein; 43 AA.
AC AAR11409;
XX
XX 25-MAR-2003 (revised)
DT 10-JUN-1991 (first entry)
XX
XX Hepatitis B surface antigen pPDI-110-148 complex epitope.
DE
XX HBSAg; chimaera; hepadnavirus; vaccine; pPV-Nhe.
KW
XX
OS / Synthetic.
XX
XX
XX
PD 10-APR-1991.
XX
XX 19-SEP-1990; 90BP-00310264.
XX
XX 19-SEP-1989; 89GB-00021172.
PR 13-AUG-1990; 90GB-00017728.
XX
XX (WELL) WELLCOME FOUND LTD.
PA (BROW/) BROWN A L.
XX
XX Brown AL, Clarke BE, Rowlands DJ;
XX

DR WPI; 1991-103861/15.
DR N-PSDB; AAQ11233.
XX
PT Particles comprising chimeric hepadnavirus core antigen protein - contg.
PT foreign aminoacid sequence comprising epitope of e.g. hepatitis A or B
PT and human immuno-deficiency virus, used as vaccine.
XX
PS Example 3; Page 10; 24pp; English.
XX
CC The insert encoding this epitope was ligated into the NheI restriction
CC site of plasmid pPV-Nhe to recreate the NheI sites which code for Ala-Ser
CC residues flanking the epitope. The recombinant plasmid was transformed
CC into E.coli strain XL-1 Blue. Positive clones were cultured to high
CC density in nutrient broth and induced to express chimaeric protein
CC comprising the desired epitope by addition of IPTG. (Updated on 25-MAR-
CC 2003 to correct PR field.) (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 43 AA;

Alignment Scores:
Pred. No.: 1.22e-10 Length: 43
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x AAR11409 (1-43)

QY 1 ACGACTCTGCTCAAGAACTCTACGTTTCCTCTTGTGCTGTACAAAACCTTCGGAC 60
Db 18 ThrThrProAlaGlnGlyAsnSerLysPheProSerCysCysThrLysProThrAsp 37

QY 61 GGAAACTGC 69
Db 38 GlyAsnCys 40

RESULT 11
ADM68797
ID ADM68797 standard; protein; 51 AA.
AC ADM68797;
XX
XX 24-MAR-2005 (first entry)
DT
XX
DE Hepatitis B virus HBSAg surface antigen fragment #7.
XX
KW antigen; HBSAg; hepatitis marker; virucide; hepatotropic;
KW antiinflammatory; DNA detection; vaccine; diagnostic;
KW hepatitis B virus infection; selectable marker.
XX
XX Hepatitis B virus.
OS
XX WO2004113369-A1.
PN
XX
PD 29-DEC-2004.
XX
PF 17-JUN-2004; 2004WO-EP006515.
XX
PR 20-JUN-2003; 2003DE-01028080.
XX
XX (DADE-) DADE BEHRING MARBURG GMBH.
XX
XX Krupka U;
PI
XX WPI; 2005-066451/07.
DR N-PSDB; ADM68786.
XX
XX New oligo- or poly-peptide variants of hepatitis B surface antigen,
PT useful for diagnosis of, and vaccination against, hepatitis infection,
PT derived from strain HDB05, also related nucleic acid and antibodies.
XX
XX Claim 3; SEQ ID NO 18; 61pp; German.
PS

XX This sequence represents novel oligo- and poly-peptides that are variants
 CC of the hepatitis B surface antigen (HBsAg). HBV strain HBD05 was isolated
 CC from a patient with inflammation of the liver but an atypical reaction
 CC pattern of hepatitis markers. PCR amplification and sequencing indicated
 CC the 5 specified mutations in the S gene, relative to adw wild-type. These
 CC mutations all caused amino acid variations: 4 in the a-determinant region
 CC and the other close to it at residue 181. The products of the invention
 CC have virucide, hepatotropic and antiinflammatory activity. The products
 CC of the invention are used to detect specific antibodies, as probes or
 CC primers for detecting hepatitis B virus (HBV) nucleic acid, as immunogens
 CC and antibodies (Ab) raised against HBsAg are used for detecting HBsAg
 CC including use in vaccination, diagnosis and screening of donated blood.
 CC The antibodies do not bind, or bind much less strongly, to the HBsAg from
 CC virus of genotype A, subtype adw. This sequence represents a fragment of
 CC Hepatitis B virus surface protein (HBsAg) corresponding to residues 115-
 XX 165.
 XX
 SQ Sequence 51 AA;

Alignment Scores:
 Pred. No.: 1.23e-10 Length: 51
 Score: 130.00 Matches: 21
 Percent Similarity: 95.65% Conservative: 1
 Best Local Similarity: 91.30% Mismatches: 1
 Query Match: 93.53% Indels: 0
 DB: 9 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x ADW68797 (1-51)

QY 1 ACAGCTCTGCTCAAGGAACCTCTACGTTTCCCTCTGTGCTGTACAAACCTTCGGAC 60
 Db 11 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProThrAsp 30
 QY 61 GGAAACTGC 69
 Db 31 GlyAsnCys 33

RESULT 12

ADW68794
 ID ADW68794 standard; protein; 55 AA.

XX
 AC ADW68794;

XX 24-MAR-2005 (first entry)

XX Hepatitis B virus HBsAg surface antigen fragment #4.

XX antigen; HBsAg; hepatitis marker; virucide; hepatotropic;
 KW antiinflammatory; DNA detection; vaccine; diagnostic;
 KW hepatitis B virus infection; selectable marker.

XX Hepatitis B virus.

XX WO2004113369-A1.

XX 29-DEC-2004.

XX 17-JUN-2004; 2004WO-EP006515.

XX 20-JUN-2003; 2003DE-01028080.

XX (DADE-) DADE BEHRING MARBURG GMBH.

XX Krupka U;

XX WPI; 2005-066451/07.

XX N-PSDB; ADW68783.

XX New oligo- or poly-peptide variants of hepatitis B surface antigen,
 PT useful for diagnosis of, and vaccination against, hepatitis infection,
 PT derived from strain HBD05, also related nucleic acid and antibodies.

PS Claim 3; SEQ ID NO 15; 61pp; German.

XX This sequence represents novel oligo- and poly-peptides that are variants
 CC of the hepatitis B surface antigen (HBsAg). HBV strain HBD05 was isolated
 CC from a patient with inflammation of the liver but an atypical reaction
 CC pattern of hepatitis markers. PCR amplification and sequencing indicated
 CC the 5 specified mutations in the S gene, relative to adw wild-type. These
 CC mutations all caused amino acid variations: 4 in the a-determinant region
 CC and the other close to it at residue 181. The products of the invention
 CC have virucide, hepatotropic and antiinflammatory activity. The products
 CC of the invention are used to detect specific antibodies, as probes or
 CC primers for detecting hepatitis B virus (HBV) nucleic acid, as immunogens
 CC and antibodies (Ab) raised against HBsAg are used for detecting HBsAg
 CC including use in vaccination, diagnosis and screening of donated blood.
 CC The antibodies do not bind, or bind much less strongly, to the HBsAg from
 CC virus of genotype A, subtype adw. This sequence represents a fragment of
 CC Hepatitis B virus surface protein (HBsAg) corresponding to residues 111-
 XX 165.
 XX

SQ Sequence 55 AA;

Alignment Scores:
 Pred. No.: 1.24e-10 Length: 55
 Score: 130.00 Matches: 21
 Percent Similarity: 95.65% Conservative: 1
 Best Local Similarity: 91.30% Mismatches: 1
 Query Match: 93.53% Indels: 0
 DB: 9 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x ADW68794 (1-55)

QY 1 ACAGCTCTGCTCAAGGAACCTCTACGTTTCCCTCTGTGCTGTACAAACCTTCGGAC 60
 Db 15 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProThrAsp 34
 QY 61 GGAAACTGC 69
 Db 35 GlyAsnCys 37

RESULT 13

ADW68793

ID ADW68793 standard; protein; 60 AA.

XX
 AC ADW68793;

XX 24-MAR-2005 (first entry)

XX Hepatitis B virus HBsAg surface antigen fragment #3.

XX antigen; HBsAg; hepatitis marker; virucide; hepatotropic;
 KW antiinflammatory; DNA detection; vaccine; diagnostic;
 KW hepatitis B virus infection; selectable marker.

XX Hepatitis B virus.

XX WO2004113369-A1.

XX 29-DEC-2004.

XX 17-JUN-2004; 2004WO-EP006515.

XX 20-JUN-2003; 2003DE-01028080.

XX (DADE-) DADE BEHRING MARBURG GMBH.

XX Krupka U;

XX WPI; 2005-066451/07.

XX N-PSDB; ADW68782.

XX New oligo- or poly-peptide variants of hepatitis B surface antigen,
 PT useful for diagnosis of, and vaccination against, hepatitis infection,
 PT derived from strain HBD05, also related nucleic acid and antibodies.

```
XX PS Claim 3; SEQ ID NO 14; 61pp; German.
XX
XX This sequence represents novel oligo- and poly-peptides that are variants
XX of the hepatitis B surface antigen (HBsAg). HBV strain HBD05 was isolated
XX from a patient with inflammation of the liver but an atypical reaction
XX pattern of hepatitis markers. PCR amplification and sequencing indicated
XX the 5 specified mutations in the S gene, relative to adw wild-type. These
XX mutations all caused amino acid variations: 4 in the a-determinant region
XX and the other close to it at residue 181. The products of the invention
XX have virucide, hepatotropic and antiinflammatory activity. The products
XX of the invention are used to detect specific antibodies, as probes or
XX primers for detecting hepatitis B virus (HBV) nucleic acid, as immunogens
XX and antibodies (Ab) raised against HBsAg are used for detecting HBsAg
XX including use in vaccination, diagnosis and screening of donated blood.
XX The antibodies do not bind, or bind much less strongly, to the HBsAg from
XX virus of genotype A, subtype adw. This sequence represents a fragment of
XX Hepatitis B virus surface protein (HBsAg) corresponding to residues 111-
XX 170.
XX SQ Sequence 60 AA;

Alignment Scores:
Pred. No.: 1.248-10 Length: 60
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0
DB: 9 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x ADM68793 (1-60)

QY 1 ACGACTCCTCGTCAAGGAACTCTACGTTTCCTCTGTGTGTACAAACCTTCGGAC 60
Db 15 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProThrAsp 34

QY 61 GGAAACTGC 69
Db 35 GlyAsnCys 37

RESULT 14
ADM68806
ID ADM68806 standard; protein; 70 AA.
XX
XX AC ADM68806;
XX
XX DT 24-MAR-2005 (first entry)
XX
XX DE Hepatitis B virus genotype A HBsAg fragment.
XX
XX KW antigen; HBsAg; hepatitis marker; virucide; hepatotropic;
XX antiinflammatory; DNA detection; vaccine; diagnostic;
XX KW hepatitis B virus infection; selectable marker; a-determinant.
XX
XX OS Hepatitis B virus.
XX
XX PN WO2004113369-A1.
XX
XX PD 29-DEC-2004.
XX
XX PF 17-JUN-2004; 2004WO-BF006515.
XX
XX PR 20-JUN-2003; 2003DB-01028080.
XX
XX PA (DADE-) DADE BEHRING MARBURG GMBH.
XX
XX PI Krupka U;
XX
XX DR WPI; 2005-066451/07.
XX
XX PT New oligo- or poly-peptide variants of hepatitis B surface antigen,
XX useful for diagnosis of, and vaccination against, hepatitis infection,
XX derived from strain HBD05, also related nucleic acid and antibodies.
```

```
XX PS Disclosure; Fig 1; 61pp; German.
XX
XX This sequence represents novel oligo- and poly-peptides that are variants
XX of the hepatitis B surface antigen (HBsAg). HBV strain HBD05 was isolated
XX from a patient with inflammation of the liver but an atypical reaction
XX pattern of hepatitis markers. PCR amplification and sequencing indicated
XX the 5 specified mutations in the S gene, relative to adw wild-type. These
XX mutations all caused amino acid variations: 4 in the a-determinant region
XX and the other close to it at residue 181. The products of the invention
XX have virucide, hepatotropic and antiinflammatory activity. The products
XX of the invention are used to detect specific antibodies, as probes or
XX primers for detecting hepatitis B virus (HBV) nucleic acid, as immunogens
XX and antibodies (Ab) raised against HBsAg are used for detecting HBsAg
XX including use in vaccination, diagnosis and screening of donated blood.
XX The antibodies do not bind, or bind much less strongly, to the HBsAg from
XX virus of genotype A, subtype adw. This sequence represents a fragment of
XX Hepatitis B virus genotype A surface protein (HBsAg) a-determinant
XX corresponding to residues 101-170.
XX SQ Sequence 70 AA;

Alignment Scores:
Pred. No.: 1.268-10 Length: 70
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0
DB: 9 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x ADM68806 (1-70)

QY 1 ACGACTCCTCGTCAAGGAACTCTACGTTTCCTCTGTGTGTACAAACCTTCGGAC 60
Db 25 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProThrAsp 44

QY 61 GGAAACTGC 69
Db 45 GlyAsnCys 47

RESULT 15
ADM68821
ID ADM68821 standard; protein; 70 AA.
XX
XX AC ADM68821;
XX
XX DT 24-MAR-2005 (first entry)
XX
XX DE Hepatitis B virus wild type HDB 05 HBsAg S protein fragment.
XX
XX KW antigen; HBsAg; hepatitis marker; virucide; hepatotropic;
XX antiinflammatory; DNA detection; vaccine; diagnostic;
XX KW hepatitis B virus infection; selectable marker; S protein.
XX
XX OS Hepatitis B virus.
XX
XX PN WO2004113369-A1.
XX
XX PD 29-DEC-2004.
XX
XX PF 17-JUN-2004; 2004WO-BP006515.
XX
XX PR 20-JUN-2003; 2003DE-01028080.
XX
XX PA (DADE-) DADE BEHRING MARBURG GMBH.
XX
XX PI Krupka U;
XX
XX DR WPI; 2005-066451/07.
XX
XX PT New oligo- or poly-peptide variants of hepatitis B surface antigen,
XX useful for diagnosis of, and vaccination against, hepatitis infection,
XX derived from strain HBD05, also related nucleic acid and antibodies.
```

XX
PS
XX

Disclosure; Fig 6; 61pp; German.

This sequence represents novel oligo- and poly-peptides that are variants of the hepatitis B surface antigen (HBsAg). HBV strain HBD05 was isolated from a patient with inflammation of the liver but an atypical reaction pattern of hepatitis markers. PCR amplification and sequencing indicated the 5 specified mutations in the S gene, relative to adw wild-type. These mutations all caused amino acid variations: 4 in the a-determinant region and the other close to it at residue 181. The products of the invention of the invention are used to detect specific antibodies, as probes or have virucide, hepatotropic and antiinflammatory activity. The products of the invention are used to detect specific antibodies, as immunogens and antibodies (Ab) raised against HBsAg are used for detecting HBsAg including use in vaccination, diagnosis and screening of donated blood. The antibodies do not bind, or bind much less strongly, to the HBsAg from virus of genotype A, subtype adw. This sequence represents a fragment of the Hepatitis B virus wild-type HDB 05 surface protein (HBsAg) S protein corresponding to residues 101-170.

XX .SQ Sequence 70 AA;

Alignment Scores:

Pred. No.:	1.26e-10	Length:	70
Score:	130.00	Matches:	21
Percent Similarity:	95.65%	Conservative:	1
Best Local Similarity:	91.30%	Mismatches:	1
Query Match:	93.53%	Indels:	0
DB:	9	Gaps:	0

US-10-761-006A-1_COPY_527_595 (1-69) x ADM68821 (1-70)

Qy	1	ACGACTCTGCTCAAGGAACCTTACCTTTCCCTCTTGTGCTGTACAAACCTTCGGAC	60
Db	25	ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProThrAsp	44
Qy	61	GGAACTGC	69
Db	45	GlyAsnCys	47

Search completed: December 27, 2005, 20:51:59
Job time : 20.262 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 27, 2005, 20:44:58 ; Search time 4.6 Seconds
(without alignments)
2480.270 Million cell updates/sec

Title: US-10-761-006A-1_COPY_527_595

Perfect score: 139
Sequence: 1 ACAGCTCTGCTCCTCAGGAAA.....AACCTTCGGACGGAACTGC 69

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p model -DEV=xlp
-O=/cgn2_1/USPO spoel_p/US10761006/runat_27122005_192901_15775/app query.fasta_1.1102
-DB=issued_Patents_AA -QPMT=fastan -SUFFIX=rai -MINMATCH=0 1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10761006 @Cgn 1 101 @runat_27122005_192901_15775 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents AA:
1: /cgn2_6/ptodata/1/iaa/5 COMB.pcp:*
2: /cgn2_6/ptodata/1/iaa/6 COMB.pcp:*
3: /cgn2_6/ptodata/1/iaa/H COMB.pcp:*
4: /cgn2_6/ptodata/1/iaa/PCRTUS COMB.pcp:*
5: /cgn2_6/ptodata/1/iaa/RE COMB.pcp:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pcp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	139	100.0	400	2	US-09-719-528A-3
2	139	100.0	400	2	US-10-209-264-3
3	133	95.7	154	2	US-09-193-104-25
4	130	93.5	24	1	US-08-776-585-6
5	130	93.5	24	2	US-09-282-826-3
6	130	93.5	154	2	US-09-193-104-10
7	130	93.5	154	2	US-09-193-104-20
8	130	93.5	154	2	US-09-193-104-21
9	130	93.5	225	6	5436139-4
10	130	93.5	226	6	5196194-21
11	130	93.5	226	6	5436139-5
12	130	93.5	281	2	US-09-721-480-3

13	130	93.5	390	2	US-09-721-480-5	Sequence 5, Appli
14	130	93.5	395	6	5196194-18	Patent No. 5196194
15	130	93.5	423	1	US-08-760-797A-1	Sequence 1, Appli
16	130	93.5	424	1	US-08-760-797A-3	Sequence 3, Appli
17	130	93.5	424	2	US-08-932-929B-1	Sequence 3, Appli
18	130	93.5	424	2	US-08-932-929B-3	Sequence 7, Appli
19	130	93.5	531	2	US-09-721-480-7	Patent No. 5196194
20	127	91.4	40	6	5196194-8	Sequence 5, Appli
21	126	90.6	92	2	US-09-193-104-5	Sequence 7, Appli
22	125	89.9	154	2	US-09-193-104-6	Sequence 6, Appli
23	124	89.2	92	2	US-09-193-104-9	Sequence 9, Appli
24	124	89.2	154	2	US-09-193-104-13	Sequence 13, Appli
25	124	89.2	154	2	US-09-193-104-14	Sequence 14, Appli
26	124	89.2	154	2	US-09-193-104-15	Sequence 15, Appli
27	124	89.2	154	2	US-09-193-104-16	Sequence 16, Appli
28	124	89.2	154	2	US-09-193-104-17	Sequence 17, Appli
29	124	89.2	154	2	US-09-306-420C-20	Sequence 20, Appli
30	124	89.2	157	2	US-09-306-420C-20	Sequence 20, Appli
31	123	88.5	226	6	5198348-1	Patent No. 5198348
32	122	87.8	154	2	US-09-193-104-18	Sequence 18, Appli
33	121	87.1	92	2	US-09-193-104-4	Sequence 4, Appli
34	121	87.1	154	2	US-09-193-104-12	Sequence 12, Appli
35	121	87.1	154	2	US-09-193-104-19	Sequence 19, Appli
36	121	87.1	154	2	US-09-193-104-22	Sequence 22, Appli
37	121	87.1	154	2	US-09-193-104-23	Sequence 23, Appli
38	121	87.1	154	2	US-09-193-104-24	Sequence 24, Appli
39	121	87.1	226	1	US-08-378-011A-3	Sequence 3, Appli
40	121	87.1	226	2	US-09-471-573A-2	Sequence 2, Appli
41	121	87.1	226	2	US-09-471-573A-40	Sequence 40, Appli
42	121	87.1	236	1	US-08-378-011A-1	Sequence 1, Appli
43	121	87.1	277	6	5164485-2	Patent No. 5164485
44	117	84.2	154	2	US-09-193-104-8	Sequence 8, Appli
45	116	83.5	228	1	US-08-447-591-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-719-528A-3
; Sequence 3, Application US/09719528A
; Patent No. 6558675
; GENERAL INFORMATION:

APPLICANT: Oon, Chong Jin
Lim, Gek Keow

Chen, Wei Ning
Zhao, Yi

TITLE OF INVENTION: A MUTANT HUMAN HEPATITIS B VIRAL STRAIN AND
USES THEREOF

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Ladass & Parry
STREET: 26 West 61 Street

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10023

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/719,528A

FILING DATE: 30-Apr-2001

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/SG98/00046

FILING DATE: 19-JAN-1998

ATTORNEY/AGENT INFORMATION:

NAME: Mass, Clifford J.

REGISTRATION NUMBER: 30,086

REFERENCE/DOCKET NUMBER: U-013109-7

TELECOMMUNICATION INFORMATION:

```
/
/ TELEPHONE: (212) 708-1800
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 400 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-719-528A-3

Alignment Scores:
Pred. No.: 1.91e-13 Length: 400
Score: 139.00 Matches: 23
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x US-09-719-528A-3 (1-400)

QY 1 ACGACTCCTGCTCAAGAACTCTACGTTTCCCTCTGTGCTGTACAAAACCTTCGGAC 60
Db 299 ThrThrProAlaGlnGlyAsnSerThrPheProSerCysCysThrLysProSerAsp 318

QY 61 GGAAGACTGC 69
Db 319 GlyAsnCys 321

RESULT 2
US-10-209-264-3
; Sequence 3, Application US/10209264
; Patent No. 6787142
; GENERAL INFORMATION:
; APPLICANT: Oon, Chong Jin
; Zhao, Yi
; Lin, Gek Keow
; Chen, Wei Ning
; TITLE OF INVENTION: A MUTANT HUMAN HEPATITIS B VIRAL STRAIN AND
; USES THEREOF
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ladas & Parry
; STREET: 26 West 61 Street
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10023
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/10/209,264
; APPLICATION NUMBER: US/10/209,264
; FILING DATE: 31-Jul-2002
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/SG98/00046
; FILING DATE: 19-JAN-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maas, Clifford J.
; REGISTRATION NUMBER: 30,086
; REFERENCE/DOCKET NUMBER: U-013109-7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 708-1800
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 400 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-209-264-3
```

```
Alignment Scores:
Pred. No.: 1.91e-13 Length: 400
Score: 139.00 Matches: 23
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x US-10-209-264-3 (1-400)

QY 1 ACGACTCCTGCTCAAGAACTCTACGTTTCCCTCTGTGCTGTACAAAACCTTCGGAC 60
Db 299 ThrThrProAlaGlnGlyAsnSerThrPheProSerCysCysThrLysProSerAsp 318

QY 61 GGAAGACTGC 69
Db 319 GlyAsnCys 321

RESULT 3
US-09-193-104-25
; Sequence 25, Application US/09193104A
; Patent No. 6172193
; GENERAL INFORMATION:
; APPLICANT: Primi, Daniele
; APPLICANT: Fiordalisi, Gianfranco
; APPLICANT: Palla, Mario
; TITLE OF INVENTION: Escape Mutant of the Surface Antigen of Hepatitis B
; TITLE OF INVENTION: Virus
; FILE REFERENCE: SBD1004US
; CURRENT APPLICATION NUMBER: US/09/193,104A
; CURRENT FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: EP 97830635.5
; EARLIER FILING DATE: 1997-12-01
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Hepatitis B virus
US-09-193-104-25

Alignment Scores:
Pred. No.: 1.53e-12 Length: 154
Score: 133.00 Matches: 22
Percent Similarity: 95.65% Conservative: 0
Best Local Similarity: 95.65% Mismatches: 1
Query Match: 95.68% Indels: 0
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x US-09-193-104-25 (1-154)

QY 1 ACGACTCCTGCTCAAGAACTCTACGTTTCCCTCTGTGCTGTACAAAACCTTCGGAC 60
Db 92 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProSerAsp 111

QY 61 GGAAGACTGC 69
Db 112 GlyAsnCys 114

RESULT 4
US-08-776-585-6
; Sequence 6, Application US/08776585
; Patent No. 5932426
; GENERAL INFORMATION:
; APPLICANT: BARALLE, Francesco E.
; APPLICANT: SCODELLER, Eduino
; APPLICANT: TISMINEZKY, Sergio
; TITLE OF INVENTION: MOLECULAR PRESENTING SYSTEM
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
```

CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20007-5109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/776,585
 FILING DATE: 31-MAR-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO PCT/EP95/03114
 FILING DATE: 04-AUG-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: ISACSON, John P.
 REGISTRATION NUMBER: 33,715
 REFERENCE/DOCKET NUMBER: 69582/106
 TELEPHONE: (202)672-5300
 TELEFAX: (202)672-5399
 TELEX: 904136
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 24 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-776-585-6

Alignment Scores:
 Pred. No.: 3,79e-12 Length: 24
 Score: 130.00 Matches: 21
 Percent Similarity: 95.65% Conservative: 1
 Best Local Similarity: 91.30% Mismatches: 1
 Query Match: 93.53% Indels: 0
 DB: 1 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x US-08-776-585-6 (1-24)

QY 1 ACGACTCTGCTCAAGAACTCTACGTTCCCTCTGTGCTGACAAACCTTCGGAC 60
 Db 2 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProThrAsp 21

QY 61 GGAAGACTGC 69
 Db 22 GlyAsnCys 24

RESULT 5
 US-09-282-826-3
 ; Sequence 3, Application US/09282826A
 ; Patent No. 6403092
 ; GENERAL INFORMATION:
 ; APPLICANT: Pizzo, Salvatore
 ; APPLICANT: Gron, Hanne
 ; TITLE OF INVENTION: IMMUNE RESPONSE MODULATOR ALPHA-2 MACROGLOBULIN COMPLEX
 ; FILE REFERENCE: 2295-1-001CIP
 ; CURRENT APPLICATION NUMBER: US/09/282,826A
 ; CURRENT FILING DATE: 1999-03-31
 ; EARLIER APPLICATION NUMBER: 09/053,301
 ; EARLIER FILING DATE: 1998-04-01
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 24
 ; TYPE: PRT
 ; ORGANISM: HBSAG
 US-09-282-826-3

Alignment Scores:
 Pred. No.: 3,79e-12 Length: 24

Score: 130.00 Matches: 21
 Percent Similarity: 95.65% Conservative: 1
 Best Local Similarity: 91.30% Mismatches: 1
 Query Match: 93.53% Indels: 0
 DB: 1 Gaps: 0
 US-10-761-006A-1_COPY_527_595 (1-69) x US-09-282-826-3 (1-24)

QY 1 ACGACTCTGCTCAAGAACTCTACGTTCCCTCTGTGCTGACAAACCTTCGGAC 60
 Db 2 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProThrAsp 21
 QY 61 GGAAGACTGC 69
 Db 22 GlyAsnCys 24

RESULT 6
 US-09-193-104-10
 ; Sequence 10, Application US/09193104A
 ; Patent No. 6172193
 ; GENERAL INFORMATION:
 ; APPLICANT: Primi, Daniele
 ; APPLICANT: Fiordalisi, Gianfranco
 ; APPLICANT: Palla, Mario
 ; TITLE OF INVENTION: Escape Mutant of the Surface Antigen of Hepatitis B
 ; TITLE OF INVENTION: Virus
 ; FILE REFERENCE: SBD1004US
 ; CURRENT APPLICATION NUMBER: US/09/193,104A
 ; CURRENT FILING DATE: 1998-11-16
 ; EARLIER APPLICATION NUMBER: EP 97830635.5
 ; EARLIER FILING DATE: 1997-12-01
 ; NUMBER OF SEQ ID NOS: 44
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 10
 ; LENGTH: 154
 ; TYPE: PRT
 ; ORGANISM: Hepatitis B virus
 US-09-193-104-10

Alignment Scores:
 Pred. No.: 4,51e-12 Length: 154
 Score: 130.00 Matches: 21
 Percent Similarity: 95.65% Conservative: 1
 Best Local Similarity: 91.30% Mismatches: 1
 Query Match: 93.53% Indels: 0
 DB: 1 Gaps: 0
 US-10-761-006A-1_COPY_527_595 (1-69) x US-09-193-104-10 (1-154)

QY 1 ACGACTCTGCTCAAGAACTCTACGTTCCCTCTGTGCTGACAAACCTTCGGAC 60
 Db 92 ThrThrProAlaGlnGlyAsnSerMetTyProSerCysCysThrLysProSerAsp 111
 QY 61 GGAAGACTGC 69
 Db 112 GlyAsnCys 114

RESULT 7
 US-09-193-104-20
 ; Sequence 20, Application US/09193104A
 ; Patent No. 6172193
 ; GENERAL INFORMATION:
 ; APPLICANT: Primi, Daniele
 ; APPLICANT: Fiordalisi, Gianfranco
 ; APPLICANT: Palla, Mario
 ; TITLE OF INVENTION: Escape Mutant of the Surface Antigen of Hepatitis B
 ; TITLE OF INVENTION: Virus
 ; FILE REFERENCE: SBD1004US
 ; CURRENT APPLICATION NUMBER: US/09/193,104A
 ; CURRENT FILING DATE: 1998-11-16
 ; EARLIER APPLICATION NUMBER: EP 97830635.5
 ; EARLIER FILING DATE: 1997-12-01
 ; NUMBER OF SEQ ID NOS: 44


```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Hepatitis B virus
US-09-193-104-20

Alignment Scores:
Pred. No.: 4,51e-12 Length: 154
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x US-09-193-104-20 (1-154)

QY 1 ACGACTCCTGCTCAAGAACTCTACGTTCCCTCTTGTGCTGTACAAAACCTTCGGAC 60
Db 92 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProThrAsp 111
QY 61 GGAAACTGC 69
Db 112 GlyAsnCys 114

RESULT 8
US-09-193-104-21
; Sequence 21, Application US/09193104A
; Patent No. 6172193
; GENERAL INFORMATION:
; APPLICANT: Primi, Daniele
; APPLICANT: Fiordalisi, Gianfranco
; APPLICANT: Palla, Mario
; TITLE OF INVENTION: Escape Mutant of the Surface Antigen of Hepatitis B
; FILE REFERENCE: SBD1004US
; CURRENT APPLICATION NUMBER: US/09/193,104A
; CURRENT FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: EP 97830635.5
; EARLIER FILING DATE: 1997-12-01
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Hepatitis B virus
US-09-193-104-21

Alignment Scores:
Pred. No.: 4,51e-12 Length: 154
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x US-09-193-104-21 (1-154)

QY 1 ACGACTCCTGCTCAAGAACTCTACGTTCCCTCTTGTGCTGTACAAAACCTTCGGAC 60
Db 92 ThrThrProAlaGlnGlyAsnSerLeuPheProSerCysCysThrLysProThrAsp 111
QY 61 GGAAACTGC 69
Db 112 GlyAsnCys 114

RESULT 9
5436139-4
; Patent No. 5436139
; APPLICANT: RUTTER, WILLIAM J.; GOODMAN, HOWARD M.
; TITLE OF INVENTION: NON-PASSAGEABLE VIRUS
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/058/89,993
; FILING DATE: 08-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 571,331
; FILING DATE: 22-AUG-1990
; APPLICATION NUMBER: 513,055
; FILING DATE: 12-JUL-1983
; APPLICATION NUMBER: 107,267
; FILING DATE: 21-DEC-1979
; APPLICATION NUMBER: 41,909
; FILING DATE: 24-MAY-1979
; SEQ ID NO: 4
; LENGTH: 225
5436139-4

Alignment Scores:
Pred. No.: 4,67e-12 Length: 225
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0
DB: 6 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x 5436139-4 (1-225)

QY 1 ACGACTCCTGCTCAAGAACTCTACGTTCCCTCTTGTGCTGTACAAAACCTTCGGAC 60
Db 124 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProThrAsp 143
QY 61 GGAAACTGC 69
Db 144 GlyAsnCys 146

RESULT 10
5196194-21
; Patent No. 5196194
; APPLICANT: RUTTER, WILLIAM J.; GOODMAN, HOWARD M.
; TITLE OF INVENTION: VACCINES CONTAINING HEPATITIS B S-PROTEIN
; NUMBER OF SEQUENCES: 21
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/679,621
; FILING DATE: 7-DEC-1984
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 513,055
; FILING DATE: 12-JUL-1983
; APPLICATION NUMBER: 107,267
; FILING DATE: 21-DEC-1979
; APPLICATION NUMBER: 41,909
; FILING DATE: 24-MAY-1979
; SEQ ID NO: 21
; LENGTH: 226
5196194-21

Alignment Scores:
Pred. No.: 4,67e-12 Length: 226
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0
DB: 6 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x 5196194-21 (1-226)

QY 1 ACGACTCCTGCTCAAGAACTCTACGTTCCCTCTTGTGCTGTACAAAACCTTCGGAC 60
Db 125 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProThrAsp 144
QY 61 GGAAACTGC 69
Db 145 GlyAsnCys 147

RESULT 11
5436139-5
```

[illegible]

DB: 6 Gaps: 0
US-10-761-006A-1_COPY_527_595 (1-69) x 5196194-18 (1-395)
QY 1 AGGACTCTGCTCAGGAAGCTAGTTTCCTCTTGCTGTACAAAACTTCGGAC 60
Db 294 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysCysThrLysProThrAsp 313
QY 61 GGAAACTGC 69
Db 314 GlyAsnCys 316

Db 342 GlyAsnCys 344

Search completed: December 27, 2005, 21:03:41
Job time : 4.6 secs

102A

RESULT 15
US-08-760-797A-1
; Sequence 1, Application US/08760797A
; Patent No. 5935902
; GENERAL INFORMATION:
; APPLICANT: De Wilde, Michel
; APPLICANT: Cohen, Joseph
; TITLE OF INVENTION: Hybrid Protein Between CS
; TITLE OF INVENTION: from Plasmodium and HBsAg
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760,797A
; FILING DATE: 04-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/442,612
; FILING DATE: 17-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baumeister, Kirk
; REGISTRATION NUMBER: 33,833
; REFERENCE/DOCKET NUMBER: B45015-1C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5096
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 423 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-760-797A-1

Alignment Scores:
Pred. No.: 4.95e-12 Length: 423
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservativeness: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0
DB: 1 Gaps: 0
US-10-761-006A-1_COPY_527_595 (1-69) x US-08-760-797A-1 (1-423)
QY 1 AGGACTCTGCTCAGGAAGCTAGTTTCCTCTTGCTGTACAAAACTTCGGAC 60
Db 322 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysCysThrLysProThrAsp 341
QY 61 GGAAACTGC 69
Db 314 GlyAsnCys 316